Universidad Autónoma de Madrid
Faculty of Science
Department of Molecular Biology

METHODS TO PREDICT ANTIBIOTIC RESISTANCE: FROM GENES TO METAGENOMES

PhD Thesis

Felipe Lira

Madrid, 2015

Universidad Autónoma de Madrid
Faculty of Science
Department of Molecular Biology

METHODS TO PREDICT ANTIBIOTIC RESISTANCE: FROM GENES TO METAGENOMES

Memory presented by Felipe Lira to obtain the
PhD degree in Molecular Biology at
Universidad Autónoma de Madrid
by:

Felipe Lira

Thesis Director
Dr. José Luis Martínez Menéndez

Thesis Tutor
Dr. Miguel Angel de Pedro

"(…) when you have eliminated the impossible, whatever remains, however improbable, must be the truth."

The Sign of the Four (1890)

Arthur Conan Doyle

Acknowledgments

I would like to express my gratitude and thanks to my Thesis supervisor, Dr. Jose Luis Martínez Menéndez, whose understanding and patience to find the results behind the data, added considerably to my experience. I appreciate his skill and vast knowledge to discuss about unusual areas. I would like to thank also to my Thesis Tutor, Dr. Miguel Angel de Pedro.

A special thanks goes out to Dr. Victor de Lorenzo and Dr. Sérgio Ricardo Nozawa.

Thank you for all administrative and technical staff who make possible all activities at the CNB.

I recognize that this research would not have been possible without the financial assistance provided by the Fundación "La Caixa" through the fellowship offered by it International PhD Program and the Universidad Autónoma de Madrid and the Spanish National Center of Biotechnology CNB – CSIC (Madrid, Spain) to support and contribute all aspects of my professional formation. In addition, the financial support by the 7th Framework Programme for Research and Technological Development through the Confronting the clinical relevance of biocide induced antibiotic resistance project (BIOHYPO) and the Evolution and Transfer of Antibiotic Resistance (EVOTAR) must be recognized to make possible to develop this Thesis.

One PhD Thesis is not a lonely journey and all the partners must be recognized for their contributions. For this reason, I extent my acknowledgments to Dra. Filiz Dilek and Dra. Ülkü Yetiş from the Environmental Engineering Department at the Middle East Technical University (Ankara, Turkey); to Francisco Fernandes and Ana Teresa Freitas from the Technical University of Lisbon INESC-ID/IST (Lisbon, Portugal); to Dr. Marco Rinaldo Oggioni from the University of Siena (Siena, Italy) and to Dr. Morten Otto Alexander Sommer from the Technical University of Denmark - DTU (Lyngby, Denmark).

I would also like to thank to my lab partners that took part at the beginning of this history and contributed directly or indirectly to the development of this work. My special thanks to Jorge Olivares (*Super Jorge*) for sharing his passion and enthusiasm for science and to teach me the Chilean language. We are still waiting you to make the "*Camino de Santiago*"; to Alicia (*La Cacereña*), always ready to provide interesting and peculiar topics to discuss after lunch, and to introduce me to the addictives *Perrunillas*; to Carolina (*La Abuela*) who always shared spontaneously her critical and analytical thinking about the human behavior; to Leonor (*La Leo*), who foresaw and warned all about my presence in the lab advising the others to take care.

Others partners also made pleasant my stay in the lab along these years and equally deserve my thankful. For this reason, I wonder to thank to Blanca for her patience and effort to teach me the basic steps in the lab; to Fernando, for helping me with advices to speak Spanish correctly. I hope you learned to speak Portuguese better than I speak Spanish; to Guillermo who demonstrated that persistence can also be the key to success. Along the way others joined in the fun to complete the party. Thank you Alejandra to convince me that

Argentines are good people, at least you; thank you Sara to inspire/suggest new modalities for the Laboratories Olympics Games as the “Extreme Research & Development”; I admire you Manu, that never gave up trying to prove that a *Cruzcampo* is better in *Sevilla* than in another part of the World. You were right! Thank you Jose Antonio for always bring positive vibrations in each visit to the lab. What to say about the newcomer, Paula? Excepting the fact that she speaks too much, take care with her. She has shown that is smart and learns fast. I need to mention a particular thank to Trini and her daily fulfilling happiness. A few doors down, I wonder to thank to Luis, always available to give an express course of “*Ingeniería de Vacaciones*”, to Sofia, because of your persistence now I believe that Sebúlcor realy exists; to Emma and her effort to keep in safe the secrets of the men; to Ruggero always auspicious and clear in his comments but still needs to accept the “*pasta con chorizo*” to be completely integrated to the Spanish culture.

I would also like to thank to my parents and brothers for the support they provided me through my entire life and in particular, to Frida for always being by my side supporting and encouraging in all circumstances. Without your partnership and patience this Thesis would be harder than it was.

Abstract

As many antibiotics exist as many mechanisms of resistance will rise. Antibiotic resistance is a worldwide problem and deserves all sort of attention and dedication to identify the critical points which might promote or facilitate the emergence of novel resistance genes in one community, as well the propagation of the already known genes. The increasing of antimicrobial resistant organisms, mediated by the transference of genes vertically or horizontally and by spontaneous mutations represent a risk for human health. Frequently, bacteria suffer injuries from antibiotics, biocides and heavy-metals, in clinical and environmental environments, which exert a selective pressure over these organisms interfering at the abundance and composition of environmental communities. This pressure of antimicrobials compounds leads to evolutionary and ecological consequences that are not fully understood and are highlighted in this work.

For this purpose, we performed a comprehensive study of the effects leaded by antibiotics and biocides from the smallest genetic elements, as genes, passing through the organized arrangement of the genome and finishing in the complex group forming the metagenome. To this, wet-lab techniques and bioinformatics approaches were applied to elucidate the main questions of this Thesis: (1) Does the presence of plasmids containing quinolone resistance genes affect the fitness of the cell? (2) Low-level resistance *qnr* genes can evolve towards high-level resistance?; (3) Infective and environmental strains of *Stenotrophomonas maltophilia* constitute two different phylogenetic branches?; (4) May the biocide triclosan exerts influence at the selection of antibiotic resistance genes and genetic mobile elements involved in the antimicrobial resistance in sludge from Waste Water Treatment (WWTP); (5) How does the presence of triclosan affects the taxonomic composition of an environmental microbiota?

The answers for our questions were: (1) In the absence of quinolone selective pressure, the plasmids containing *qnr* resistance genes are expulsed from the cell to minimize the costs of maintaining useless elements inside; (2) Cells with the capability of generating their own mechanisms of resistance and adaptation to antibiotics do not make use of foreign helper genes and mutate their genomes; (3) Independently on the origin, clinical and environmental strains of *S. maltophilia* present similar genomic composition; there is not an "infective" *S. maltophilia* lineage; (4) Triclosan alters the number of antibiotic resistance genes present in a microbiota in a concentration-dependent way; (5) Triclosan affects the composition of the microbiota collected from a WWTP.

Altogether, these results serve to understand the multi-level procedure that operates for selecting antibiotic resistance and propose different complimentary approaches to study this problem.

Presentación

Cuantos más antibióticos existan más mecanismos de resistencia surgirán. La resistencia antimicrobiana es un problema mundial que requiere el desarrollo de los esfuerzos necesarios identificar los puntos críticos que pueden promover o facilitar la emergencia de nuevos genes de resistencia en una comunidad, así como la propagación de genes ya conocidos. El aumento del número de organismos resistentes a los antimicrobianos causado tanto por la transferencia vertical y/o horizontal de genes como por mutaciones espontáneas representa un riesgo para la salud humana. Las bacterias sufren con frecuencia el daño producido por los antibióticos, biocidas y metales pesados, tanto en hábitats clínicos como ambientales, lo que produce una presión selectiva sobre los organismos, que a su vez puede alterar la abundancia y la composición de las comunidades ambientales. Esta presión selectiva causada por los compuestos antimicrobianos puede tener consecuencias evolutivas y ecológicas que no son completamente comprendidas y que son resaltadas en este trabajo.

A tal fin, se realizó un estudio exhaustivo de los efectos causados por los antibióticos y biocidas desde los elementos genéticos más pequeños, como son los genes, pasando por la disposición y organización del genoma hasta elementos génicos más complejos como son los metagenomas. Para esto, utlozamos, tanto estudios experimentales, como aproximaciones en Bioinformática para analizar las principales preguntas de esta tesis: (1) ¿La presencia de plásmidos conteniendo los genes de resistencia a quinolonas afectan a la capacidad competitiva (*fitness*) de la célula? (2) ¿Los genes de resistencia a quinolonas de bajo nivel pueden evolucionar hacia la resistencia de alto nivel?; (3) ¿Las cepas de *Stenotrophomonas maltophilia* clínicas y ambientales constituyen dos ramas filogenéticas distintas?; (4) ¿El biocida triclosan ejerce influencia en la selección de genes de resistencia a antibióticos y elementos genéticos móviles que intervienen en la resistencia a los antimicrobianos en los lodos de las Estaciones de Tratamiento de Aguas Residuales?; (5) ¿Cómo la presencia de Triclosan afecta la composición taxonómica de la microbiota del medio ambiente?

Las respuestas para nuestras preguntas fueron: (1) En ausencia de presión selectiva por quinolona los plásmidos que contienen genes de resistencia para este antibiótico son expulsados del interior de la célula minimizando los costes de mantenimiento de elementos poco útiles presentes en su interior; (2) Las células con la capacidad de generar sus propios mecanismos de resistencia y adaptación a los antibióticos no hacen uso de genes externos, sino que se seleccionan mutaciones en sus genomas; (3) Independientemente del origen, las cepas clínicas y ambientales de *Stenotrophomonas maltophilia* presentan una composición genómica semejantes no existiendo un linaje "infeccioso" de *S. maltophilia*; (4) El número de genes de resistencia a antibióticos en una microbiota es dependiente de la concentración de triclosan; (5) El biocida triclosan afecta a la composición de la microbiota de muestras colectadas de Estaciones de Tratamiento de Aguas Residuales.

En conjunto, estos resultados sirven para entender los diferentes niveles de organización que actúan en la selección de resistencia a los antibióticos y proponer diferentes enfoques complementarios para estudiar este problema.

Table of contents

List of figure

List of table

List of Appendix

Introduction

"*The speaker who hops to the platform, skips the introduction, and jumps to the conclusion is roundly applauded*."
Unknown

1. Introduction

The use and abuse of antibiotics and biocides has evolutionary and ecological consequences that, despite several decades of research on this topic, are not fully understood. Soon after the introduction of antibiotics, it was evident that bacteria were able to evolve and acquire resistance genes. Indeed, antibiotic resistance driven by mutations and Horizontal Gene Transfer (HGT) of genes that confer antibiotic resistance is a common event in microbial populations. This shift (from susceptible to resistant) may also alter the overall composition of the microbiota promoting a second order effect on the microbial world guided by the selective forces imposed by antibiotics usage (Gillings, 2013). Among the consequences attributed to the abuse of xenobiotics it is possible to cite the appearance of novel resistance plasmids, the fixation of integrons into populations or the establishment of genomic islands at high frequency in diverse cell lineages among others. There is evidence that antibiotic resistance is a worldwide problem with serious consequences for humans, in which the interplay between ecology, evolution and natural selection must be taken into consideration (Palumbi, 2001)

It is known that a more prudent use of antimicrobials, particularly in the treatment of human disease, but also in veterinary practice, animal husbandry and agriculture, could make a significant impact on the pace and extent to which resistance emerges in microorganisms pathogenic to man. It is known that many antibiotics used to prevent and/or treat microbial infections in veterinary medicine are structural relatives to those used for treating human infections and resistance to one of them can also confer cross-resistance towards the human-used antibiotics (Berman and Riley, 2013; Mole, 2013). Nevertheless, and although the mechanisms by which organisms acquire resistance are often well understood, including the selective pressures arising from exposure to antimicrobials, the evolutionary forces and bottlenecks driving the emergence, spread and fixation of drug resistance have not been elucidated in full. In particular the role that non-clinical, natural ecosystems may have in the origin, evolution and spread of antibiotic resistance, is a topic that did not receive more attention before, but is now the subject of several studies (Angulo et al., 2004; Davies and Davies, 2010; Martínez et al., 2011).

To be effective, the control and prevention of infection due to resistant microorganisms must be an integral part of the prevention and management of communicable diseases in general. Thus, describing the distribution of infection due to resistant organisms within populations, together with changes in patterns of those infections over time, provides the basic information for taking actions to both control disease caused by resistant microorganisms and to contain the emergence of resistance. Used in conjunction with disease prevention and infection control procedures and data on antibiotic usage, strategies can be developed to protect the public health now and in the future. It is important to remind that antibiotic resistance is not a cause but the consequence of the human behavior, which forgets the importance of considering that

solving a health problem today might affect the environment, as well as human health, tomorrow. In this regards it is important to mention that most studies on antibiotic resistance focus on the analysis of bacterial pathogens, mainly at clinical settings. However, resistance genes have been acquired from environmental bacteria and non-clinical (natural ecosystems) constitute a large reservoir of resistance determinants. In this regard, studies addressing, the mechanisms of resistance used by bacteria as well as the evolutionary forces involved in the acquisition and spread of antibiotic resistance are needed for making possible to predict "future steps" in such development (Fajardo et al., 2009; Martínez et al., 2007; Stokes and Gillings, 2011).

Antimicrobial resistance: a natural problem with a potential synthetic solution

Bacteria that predate humans have been in contact with other microorganisms by billions of years and have evolved a complex series of mechanisms to survive under extreme conditions and in the presence of numerous toxic metabolites, antibiotics included. Alexander Fleming in 1928 described the first natural antibiotic minimizing the negative effects exerted by the bacteria influencing the human life. This opened the challenge to find natural products or compounds to defeat infectious microorganisms until present days (World Health Organization, 2014). By naturally-produced antibiotics we mean genetically encoded small molecules that have evolved though natural selection, which interact with biological macromolecules, as proteins, nucleic acids, carbohydrates and membranes (Wright, 2014).

Bioactive secondary metabolites, commonly used as antibiotics, may be produced by bacteria to provide a competitive growth advantage by killing susceptible neighbors in the environment (Clardy and Walsh, 2004). Because of this, natural bacteria (including antibiotic producers) must have systems to avoid the action of antibiotics. The constant use of these compounds, has selected for the emergence of resistant organisms that may be rare in the initial population but become increasingly prevalent under selective drug pressure. The pre-existence in nature of antibiotic resistance genes has favor this transfer, In addition of being the evolution force selecting for antibiotic resistance, the presence of an antibiotic can accelerate mutation and recombination in bacterial populations, leading to resistance and hence contributing to its own obsolescence (Cirz et al., 2005).

However, discovering antibiotics is not so easy and molecules have shown up in screening assays designed to find compounds to kill bacteria. This suggests that several microorganisms must be screened to identify substantially different chemical scaffolds. In the past, microorganisms were genetically modified by chemically induced mutations to obtain strains presenting increased productivity of desired metabolites (Wang et al., 2012). Today, a new families of drugs are made semi-synthetically derived from the structures of natural products, or completely synthetically with the aim of reducing the chances of selecting natural pre-existing resistance genes (Demain, 2009).

The reasoning behind was that, since these molecules were not present in Nature, bacteria should not have developed systems (antibiotic resistance genes) for avoiding their activity. At the 60's a class of antibiotics which interfere targeting the DNA gyrase (topoisomerase II) and topoisomerase IV trapping these enzymes at the DNA cleavage stage and preventing strand rejoining were created (Figure 1) Nevertheless, the broad-spectrum activity and widespread use of fluoroquinolones has contributed to the rapid emergence of resistance. In fact, the fluoroquinolone ciprofloxacin is the most consumed antibacterial agent worldwide and have been used to treat and prevent infections caused by organisms as diverse as *Salmonella enterica*, *Campylobacter* spp., *Escherichia coli*, *Klebsiella* spp., *Pseudomonas aeruginosa*, *Neisseria gonorrhoeae* and *Streptococcus pneumonia* among others (Engberg et al., 2001; Knapp et al., 1997; Mølbak et al., 2002; Redgrave et al., 2014; Reen et al., 2011). In veterinary practice, fluoroquinolones are also very extensively used for both therapeutic and non-therapeutic purposes (Brown, 1996).

By non-therapeutic purpose we mean the fact that antimicrobials are not used exclusively to treat animals' infections, but also for increasing their weight. Multiple scientific studies have confirmed that the use of antibiotics in animals contributes to the development of resistant bacterial infections in humans (World Health Organization, 2002). Many of the antibiotics used on animals are identical or closely related to those used to prevent infections in humans, including tetracyclines, macrolides, bacitracin, penicillins, sulfonamides and in particular fluoroquinolones. As above stated, the common use of these antimicrobials in animals enable bacteria to develop antibiotic resistance even when exposed to low doses of drugs over a long period of time, contributing to the rise of pathogens that are able to defeat our shared (between animals and humans) antibiotic arsenal. Since the discovery of the growth-promoting and disease-fighting capabilities of antibiotics, farmers, fish-farmers, and livestock producers have used antibiotics in everything from apples production to aquaculture. This ongoing and often low-level dosing for growth and prophylaxis inevitably results in the development of resistance in bacteria in or near livestock, and also heightens fear of new resistant strains "jumping" between species (Cash and Narasimhan, 2000).

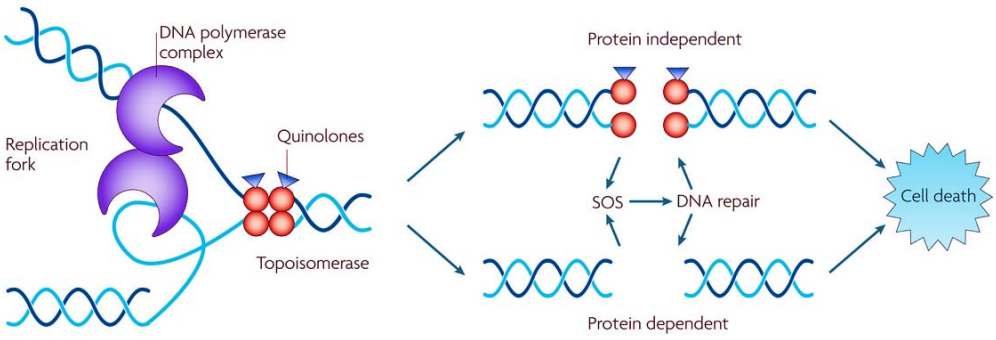


Figure 1: Quinolone antibiotics interfering with changes in DNA supercoiling by binding to topoisomerase II or IV. This leads to the formation of double-stranded DNA breaks and cell death in either a protein synthesis dependent or protein synthesis independent fashion (Figure extracted from Kohanski et al., 2010).

Upon the implementation of the quinolones as a synthetic drug, they became one of the most widely used groups of antibacterial drugs for treating human infections, but also at fish farms. As expected, mutants at the topoisomerases genes, presenting resistance to quinolones, were selected and resistance was vertically spread (Redgrave et al., 2014; Strahilevitz and Jacoby, 2009). Nevertheless, despite the synthetic origin of quinolones, plasmid encoded low level quinolone resistance has also been described (Strahilevitz and Jacoby, 2009). The most relevant plasmid-encoded quinolone resistance genes are those encoding Qnr proteins (Yang et al., 2008). These elements belong to the pentapeptide repeat family of proteins and are able to bind bacterial topoisomerases, protecting them from the action of quinolones (Drlica et al., 2009). The most widely spread *qnr* gene is *qnrA* (Sánchez and Martínez, 2012), and it has been described that the origin of this gene in in the aquatic bacterium *Shewanella algae* (Lascols et al., 2008). Whether or not the use of quinolones in fish farming has favored the jumping of *qnrA* from the *S. algae* chromosome to the plasmids of bacterial pathogens remains to be established, however, this situation indicates that any organism can be a source of resistance genes. Despite it is well established their role in low-level resistance, the role that *qnr* elements may have in the development of high level resistance to quinolones remains to be established.

The hotspots of antimicrobial resistance

It is evident that the main hotspots for the evolution and spread of antibiotics are the places where they are more widely used, the hospitals and, in a lower level, the farms. However, the constant discharge of antibiotics at natural ecosystems, may also be relevant for the selection of resistance. The human population has undergone a shift from rural to urban living, corresponding, in 1900, to 10% of the global population. In 2008, this percentage exceeded to 50% and the estimation indicates that this value will reach to 60% in the next 10 years (Grimm et al., 2008). This increase in population number, generate preoccupation because the increasingly presence of containments in streams and rivers with the variety of personal care products and antibacterial products as therapeutic drugs may select for resistance organisms at environmental settings, an issue that has received increased attention in the last years.

Most of the quantity of antibiotics used for treatment are released form the patients trough feces or urine. Consequently, undesirable antibiotic residues are released and might remain in diverse locations including non-clinical natural ecosystems (Xu et al., 2015). Hence, in addition on their direct role in selecting resistance at the point of utilization, the intensive use of antibiotics for human, veterinary and agricultural purposes, results in their continuous release into the environment, increasing the landscape for the selection of resistance. Indeed, the main concern for the release of antibiotics into the environment is connected to the development of antibiotic resistance

reducing the therapeutic potential against human and animal pathogens (Rizzo et al., 2013).

Concerning to this, many studies have revealed that sewage treatment plants and intestinal tract of animals and humans are reservoirs of drug resistance genes. In addition, several of them are identified from feces, which can commonly be used as fertilizers, that can contribute to the spread of microorganisms resistant to antimicrobials in agriculture-linked ecosystems (Penders et al., 2013; Zhou et al., 2012). In this case, the dissemination of these new resistant microorganisms into the environment may become a risk because most antibiotics and their secondary compounds are neither completely metabolized nor eliminated during sewage purification. The release of these compounds and their corresponding resistance genes into the environment might be a critical problem in the dissemination of resistance because it may promote the permanence of infectious and resistant organisms that proliferate among humans, plants and animals (Allen, 2014; Allen et al., 2010).

Recently, much more attention has focused on the role of the environment and interconnected ecological habitats, water bodies such as rivers, streams, waste water effluents, and lakes, which have been suggested to be important in facilitating the transport and transfer of the antibiotic resistance genes. Residues from hospitals, urban environment, industries, farms, aquacultures, and agriculture where the usage of antibiotics impact on the selection of resistant bacteria and promote the gene exchange are released in these ecosystems (Aminov and Mackie, 2007). Effluents from urban wastewater treatment plants (WWTPs) are main anthropogenic sources for antibiotics, antimicrobial resistance genes and antimicrobial resistant organisms.

However, the WWTP are not designed to fully remove all sort of contaminants. Recently pharmaceutical chemicals have been detected in surface waters receiving wastewater effluent in highly urbanized watersheds (Kolpin et al., 2002). As the consequence, this excessive amount of antimicrobials discharged in these treatment plants creates a propitious niche, for the selection of resistant organisms, with sub-inhibitory concentrations of antimicrobials, which can select high-level resistance to antibiotics (Farias et al., 2015; Kümmerer, 2004). In this regard, the presence of antibiotics in sewage selects for resistances that are able to spread through the microbial community and as a result, antibiotic-resistant bacteria can potentially disseminate their resistance genes widely among members of the endogenous microbial community. The sludge products of urban and rural wastewater treatment plants are increasingly used to fertilize agricultural crops, dispersing unknown amounts of resistance genes and antibiotics that withstand standard sewage treatment.

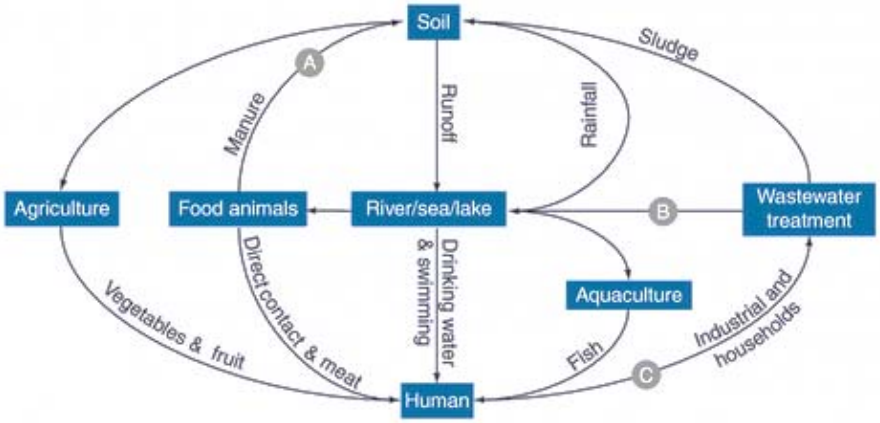


Figure 2: Potential antibiotic resistance gene dissemination. The arrows indicate possible points of dissemination among different environments where the presence of antimicrobial resistance genes were identified by metagenomic studies: (A), (B) and (C) (Figure extracted from Schmieder and Edwards 2012).

Despite the ephemerid of the transient contact between organisms present in these polluted environment, this mixing of strains and species can result in a risk for the public human health. The intense contact of organisms favor the exchange of antibiotic resistance genes mediated by the exchange of mobile elements enhancing part of the so-called resistome of one specie representing a reservoir of drug resistance mechanisms. Occurring, these bacteria may act as donor for human related bacteria turning it resistant to clinical treatments (Wright, 2010).

Induced by the idea to "protect" ourselves from infectious microorganisms, the increasing use of biocides in different circumstances may favor the establishment of mechanisms to avoid this selective pressure. In many cases, antimicrobials and biocides used in hospital care units or household products may cross the limits of a controlled sites reaching natural environments being relevant selective pressures over the environmental microorganisms. The use of molecular biology techniques applied to identify specific DNA targets without prior cultivation of their bacterial hosts has demonstrated the presence of resistance genes in clinical and municipal wastewater systems, which highlight the necessity for strategies of water quality improvement (Lupo et al., 2012). The presence of bacterial strains carrying different resistance genes found in hospital effluents containing residual antibiotics highlight their importance as a source of already resistant species that evolve under medical treatments. These organisms discharged in effluents collaborate to the widespread of antimicrobial resistant bacteria in the environment.

In spite of the great number of plasmids of different sizes containing antimicrobial antibiotic resistance genes present in the environment, and the risk these plasmids have to transfer resistance genes to other bacteria, most works on antibiotic resistance in natural ecosystems are based just in the detection of the resistance genes, without further insights on the plasmids potentially carrying them. It is then necessary to understand the mechanisms that collaborate in the wide range distribution of these plasmid-encoded resistance genes in the environment and how this affects the probabilities to occur an outbreak (Amos et al., 2014; Martínez, 2011; Petersen et al., 2002). Studies on plasmid

composition, mechanisms of co- and cross-selection and analysis on fitness costs, including the maintenance of plasmids during subculture (Andersson and Hughes, 2011; Smith and Bidochka, 1998) are required to address this important issue.

As above mentioned, most studies on antibiotic resistance focus on the selection of resistance genes. However, and although it is known that antimicrobial products and biocides are extensively used and there is mounting evidence documenting their widespread distribution in aquatic ecosystems, relatively little is known about their ecological effects on the composition of the microbiota present in these ecosystems (Cunningham et al., 2006).

Triclosan and antibiotic resistance

Triclosan [5-chloro-2-(2,4-dichlorophenoxy) phenol] (TCS) is a broad-spectrum, synthetic biocide that is active against both gram-positive and gram-negative bacteria. Currently, this biocide is incorporated into numerous consumer products as soaps, detergents, cleansers, toothpastes, surgical scrubs, hand lotions, fabrics, plastics, toothbrush handles, cutting boards, pizza-cutter, bed table tops and deodorants (Drury et al., 2013; Schweizer, 2001). Component of a wide range of products, triclosan can be introduced to natural aquatic environments mainly via domestic wastewater (Drury et al., 2013).The release of triclosan to the environment has raised concerns about resistance. Indeed, it has been demonstrated that *Escherichia coli* exposed to triclosan in laboratory conditions developed resistance suggesting that the release of triclosan to the environment could drive the development and spread of resistance among bacteria and could make triclosan less useful as an antibacterial agent, but also contribute to the selection of antibiotic resistant microorganisms (McMurry et al., 1998; Schweizer, 2001). Bacteria exposed to triclosan in laboratory conditions develop resistance through mutations in the gene *fab*I that encodes the target enzyme, as well as through overexpression of multidrug efflux pumps, the later type of mutants being antibiotic resistant (see below). It is worth mentioning that while several studies have demonstrated the ability of cultured bacteria to develop triclosan resistance based on triclosan exposure in the laboratory, few works study whether biocides can select for antimicrobial resistance, as well as address the effects of these substances in an environmental scenario (Aiello et al., 2005).

Once drained from the house, triclosan follows the same route reaching the WWTPs. These treatment stations are able to remove the majority of triclosan but not remark it completely, making possible the presence of triclosan into aquatic ecosystems in low concentrations (Bester, 2005). The problem is that triclosan is a lipophilic compound with low aqueous solubility and some studies have detected triclosan in sediments cores from 1946 (Cantwell et al., 2010). Under estuarine conditions, triclosan would accumulate in sediments presenting a half-life estimated in 540 days (Halden et al., 2006).

Due to its antimicrobial properties, triclosan may negatively affect the abundance and activity of benthic bacteria, which could have broader ecosystem-level implications. If bacterial taxa differ in triclosan sensitivity, and this compound acts as a selective agent and this situation will drive changes in bacterial community composition.

Therefore, accumulation of triclosan in sediments could increase the prevalence of resistant organisms among resident bacterial communities (Drury et al., 2013). As above stated, a disturbing aspect of triclosan release to the environment is that exposure to triclosan may select for bacteria with increased resistance not only to triclosan itself but also to other therapeutically useful antibiotics by cross-resistance (Schweizer, 2001), mainly through the overexpression bacterial multidrug efflux pumps that can confer resistance to both antibiotics and triclosan (McMurry et al., 1998).

Efflux pumps are ubiquitous in bacteria, and in many cases these efflux pumps have relatively low specificity. For example, the *AcrAB*-TolC efflux pump found in *E. coli* can pump out a variety of compounds including certain antibiotics, bile salts, cationic surfactants such as quaternary ammonium compounds, and pine oil (Ledder et al., 2006). In the case of efflux pumps extruding triclosan, exposure to this biocide in the environment could increase the frequency and spread of organisms overexpressing efflux pumps which could confer resistance to triclosan and to other antibiotics. Experimental evidence supports the connection between triclosan and antibiotic resistance. Indeed, triclosan resistant mutants, which overexpress efflux pumps, present increased resistance to antibiotics in *Pseudomonas aeruginosa* (Chuanchuen et al., 2003), *E. coli* (Braoudaki and Hilton, 2004a) and *Salmonella enterica* (Braoudaki and Hilton, 2004b) among others.

Since there exists the possibility that bacteria exposed to triclosan in the environment might develop resistance to it and that this resistance might in turn lead to the emergence and proliferation of antibiotic resistant bacteria, this could become a health concern by making some diseases more difficult to treat.

The hierarchical structure of genomes

The use of metagenomics tools is becoming the guide to the new renaissance challenge allowing a deeper knowledge on the environmental microbial communities and their components. Since 1987, when emerged the interest to extract the microbial environmental DNA directly from sediments (Ogram et al., 1987), to current days, the application of molecular approaches, as 16S rRNA gene sequencing and whole DNA sequencing, has provided unprecedented insights into the genetic dynamics within microbial communities and diversity.

When embarked in the effort to avoid culturing environmental microbial organisms, DNA fragments containing genes are treated as an unique genomic unit (Handelsman and Tiedje, 2007). The concept of *Metagenome*, as the pool of total DNA from all living organisms present in one single natural sample, was born. Diving into the deep of these data, these novel techniques provided information not accessed before,

changing the paradigms not only of the way of understanding the composition and dynamics of microbial communities but also of organisms closely related, which could only be studied before using specific genome sequencing (Tettelin et al., 2008; Wilmes et al., 2009).

Variations found between related microbial strains identified by genome sequencing were found soon after the implementation of genome-wide sequencing techniques (Tettelin et al., 2005). Nevertheless, the use of new-age genomic techniques applied to the analysis of a large number of individuals at once has allowed to identity diversity at a natural scale (Venter et al., 2004). In their seminal work, Venter and colleagues applied the 'whole-genome shotgun sequencing technique' for sequencing environmental samples improving our understanding about the complexity and variability of uncultured bacteria. In this regard, it was important to understand the flux, composition and behavior of all genomic material existent in all kind of natural habitats, how these organism may behave and how they can work together (Wilmes et al., 2009).

Naming as metagenome the collective of genomes present in one sample, all unicellular organisms cells present in a given ecosystem are named as the microbiome (Gillings, 2013; Turnbaugh et al., 2007). Going deeper in the microbiome, several genes can be shared among different species or among strains of a given bacterial species. For the later the term pangenome includes the overall genes present in a given bacterial species (Medini et al., 2005; Tettelin et al., 2008) and it is composed by two components: 1) consisting in the consensus genome formed by the whole set of genes shared by all organisms present in one clade, which is named as the core genome (van Tonder et al., 2014); and 2) the accessory genome, that encompasses genes found in some strains but not in others reflecting their unique physiological properties, including antibiotic resistance and virulence determinants in the case of bacterial pathogens (Ahmed et al., 2008; Wren, 2000).

These genes contribute to the species diversity and are not essential for growth in food-rich habitats as laboratory growth culture media. However it may confer selective advantages to occupy different niches where exist nutrients restrictions, including to those that involve avoiding the effects of the antibiotics or the colonization of the host (Jackson et al., 2011; Medini et al., 2005).

This is the case of bacterial pathogens; they may exhibit significant variation in their accessory genome reflecting the diversity of strategies that have evolved to infect their host organisms and to survive under selective pressure (Medini et al., 2005; Vernikos et al., 2015). This group of genes constitute a primary force for genome evolution and maintenance of some pathogenic bacteria because of their diverse genes related with specific phenotypes that are advantageous under certain selective conditions (Kung et al., 2010). There are cases where gene sets are specifically linked to a particular niche promoting gains to the genetic complexity and ecological plasticity (Forsman, 2014). The presence of these elements in a reduced set of individuals or in one clade is considered an

accessory genetic pool that could be acquired as a single event or by the frequent exposure to other microorganisms providing opportunities to interchange genomic material as Mobile Genetic Elements (MGE).

These elements specialized to mobilize and transfer DNA content as a mosaic structure to genomes, contribute to the generation of diversity by recombination of them. They include plasmids, integrons, insertion sequences, conjugative elements and transposons, constituting the "mobilome" and are able to carry accessory genes that may contribute with different functions to avoid the effects of antimicrobials. To note here that resistance genes may be present at both the pangenome and at the core genome (Leplae et al., 2010, 2004; Siefert, 2009a, 2009b). These resistance genes present at the pangenome form the 'resistome' and consist the group of genes capable to deal with many classes of antimicrobials (Wright, 2010). When present at the core genome, these genes are responsible for the specific phenotype of susceptibility to antibiotics of a given bacterial species and form the "intrinsic resistome" (Fajardo et al., 2009).

Organisms presenting common core genomes (Alcaraz, 2014; Mira et al., 2010) could differ in their accessory genes, and may reflect their unique physiological and virulence properties (Ahmed et al., 2008; Wren, 2000). Although, not all genetic variations are essential for adaptation and some dispensable genes may be responsible only for conferring advantages to the pathogen to persist in its host. Many authors converge with the idea that the acquisition of these genes by HGT suggests a gain and loss of genes through mutations and gene transfers, which have occurred independently in different lineages contributing to the variations of the pathogenic potentials of different organisms derived from the same ancestor (Dobrindt and Hacker, 2001; Lawrence et al., 2002).

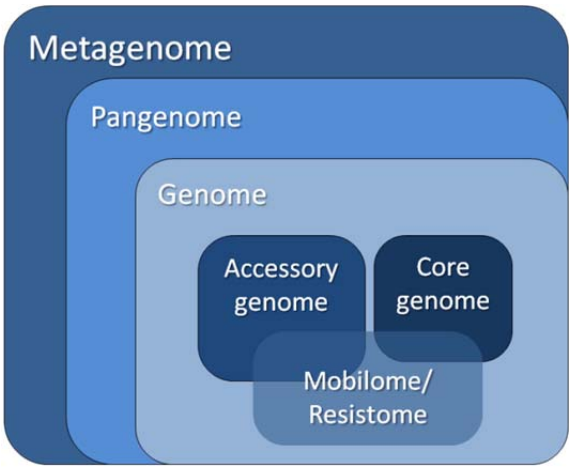


Figure 3: Spatial organization of the genomic material according the organism composition.

Therefore, a knowing how closely related strains differ or share their genomic contents (the pangenome of this species) despite being isolated from different environments can contribute to guide health programs focused in combat this opportunistic pathogens based on the plasticity presented by these organisms. In this regard, the study of the pangenome has important consequences in the way for understanding the bacterial adaptation and evolution as well their population structure (Alcaraz, 2014; Mira et al., 2010).

It is important to highlight that the hierarchical structures of bacterial genomes also provides different hierarchical levels of selection; which operates at the level of resistance gene (which also can evolve to reach higher resistance levels), but as a consequence of this selection a particular set of mobile elements, each one presenting a

hierarchical structure (see below), will be maintained, as well as the specific bacterial clones carrying these elements. Understanding the interactions among these multiple layers of selection is needed to fully understand the process of emergence and spread of antibiotic resistance (Baquero, 2004).

How and why the gene transference occurs?

Some factors constitute the ideal scenario for gene transfer. However, even when such transfer occurs, other elements are relevant for a successful fixation of the transfer event into the population. The correct deliver of the DNA sequence from the donor to the recipient and the incorporation of the acquired genes into the cell chromosome or any autonomous replicating element such as plasmids, and the expression of these genes in a significant level in the new environment constitute an important issue to the efficient transference. The relevance of the new genes for the adaptation of the bacterial recipient to the ecosystem, will allow the fixation of the transfer. Bacteria exploit three principal mechanisms for interspecies transfer of genetic elements (Dutta and Pan, 2002; Furuya and Lowy, 2006): (a) transformation, (b) transduction and (c) conjugation (Figure 4).

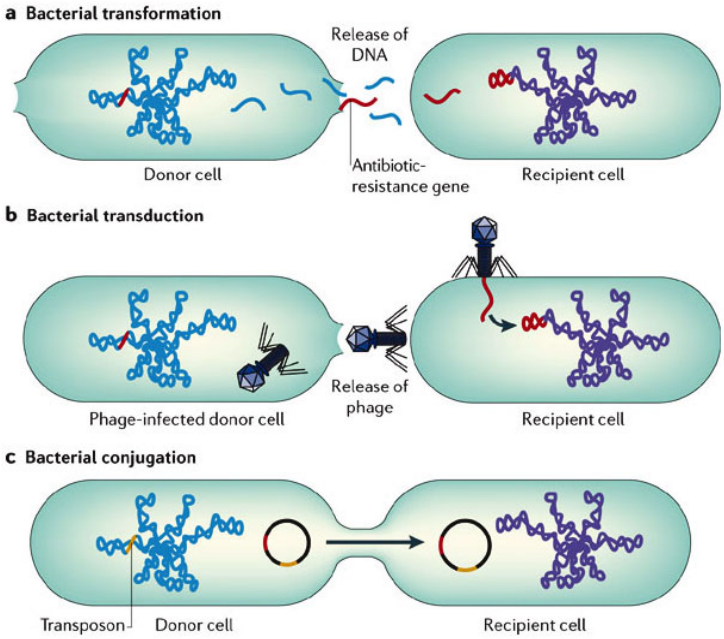


Figure 4: (a) Transformation - occurs when a single DNA fragment is released of an organism on lysis and is incorporated by another organism. The related genes can be integrated into the chromosome or plasmid of the recipient cell; (b) Transduction - when antibiotic- resistance genes are transferred from one bacterium to another under the control of bacteriophages and can be integrated into the chromosome of the recipient cell; (c) Conjugation - direct contact between two bacteria form a bridge across the bacteria which is used to exchange DNA resulting in the acquisition of antibiotic-resistance genes by the recipient cell (Furuya and Lowy, 2006).

The impact of interspecies gene transfer is radically different from that of spontaneous mutation. Point mutations can lead to the subtle refinement and alteration of existing metabolic functions, but horizontal gene transfer is a faster and wider generator of diversity (Dutta and Pan, 2002) because it has the capability of introducing, immediately upon integration, completely novel physiological traits, including those with

relevance for infection, as discussed below: (a) *Antibiotic resistance* – Bacterial organisms may acquire antibiotic resistance genes thought horizontal transfer associated whit mobile genetic elements as plasmids, integrons and transposons, that transfer genes between bacterial genomes; (b) *Pathogenicity islands* - Correspond to DNA zones which arbor virulence genes in discrete gene clusters, usually referred to as "virulence cassettes". The sequences flanking such islands frequently include short direct repeats, reminiscent of those generated upon integration of mobile genetic elements, while ORFs in certain pathogenicity islands exhibit sequence similarity to bacteriophages integrases (Lawrence et al., 2002).

Genetic mobile elements are resistance dealers

All bacteria are able to acquire resistance, either by genetic exchange or by spontaneous mutation, improving their ability to resist to one or more antimicrobials. Although mutational resistance has primary clinical importance in certain bacterial species or antibiotics, especially in the case of synthetic agents such as fluoroquinolones and oxazolidinones (Baquirin and Barlow, 2008; Woodford and Ellington, 2007), horizontal gene transfer among bacterial strains or species are considered to be the most efficient mediator of antibiotic resistance. Natural microbial ecosystems contain a large pool of potential resistance genes, that can be acquired by pathogens, together with vectors, such as plasmids, that move resistance genes between bacterial lineages, capturing many of these genes since the advent of the antibiotic era. Indeed, bacteria isolated before the medical use of antibiotics have similar plasmid backbones, but without resistance determinants, supporting that these genes were incorporated into plasmids after the human use of antibiotics. Taking into account the modularity of the Genetic Mobile Elements (GME), when these modules are captured by Lateral Gene Transfer (LGT vectors), they take with themselves the genes that compound these elements supporting the existence of multiple levels of selection (see above).

When introduced by each transfer event, the specific set of genes into a bacterial lineage remains restricted to their offspring and is absent from closely related taxa, thereby producing a scattered phylogenetic distribution. In some cases, it is possible to establish the evolutionary history of gene alleles by analyzing their distribution among various lineages. If an allele is confined to one taxon or species, being present in the same mobile element, it is more likely to have been acquired in a single gene transfer event than to have been lost independently from multiple lineages in which it was present before (Gillings and Stokes, 2012).

The modular structure and hierarchical acquisition of mobile elements is summarized in Figure 5, where the successive transfer events of GME results in a resistant plasmid vector carrying the new sequences and their "genetic firm" indicating the previous steps until to be incorporated in the "new genome". Each of these semi-autonomous DNA segment, occasionally fixed into the chromosome or present in a GME, may carry

resistance genes readily transferred between bacteria resulting in the immediate acquisition of antibiotic resistance by the recipient strain.

The transmission of GMEs between organisms occurs in two ways: once the GME has been acquired by one organism, Vertical Gene Transfer (VGT) from the ancestral strain from parental to descendent cells occurs; in addition secondary HGT events result in the acquisition of external and structured genomic material mediated via GME by the hierarchical acquisition of gene cassettes, integrons, transposons or plasmids. Within this situation, genes can spread, not only among pathogens, but among environmental microorganisms and commensals (Baquero, 2004; Stokes and Gillings, 2011). As above stated, for an effective transfer (and fixation) of this genetic material across organisms to happen some steps need to occur successfully: efficient transference of the DNA sequence from the donor into the recipient; sequence remains stable into the recipient strain, expression of the acquired genes is achieved in the new environment and the expressed element confers an adaptive advantage to the recipient host without presenting high fitness costs (Dutta and Pan, 2002).

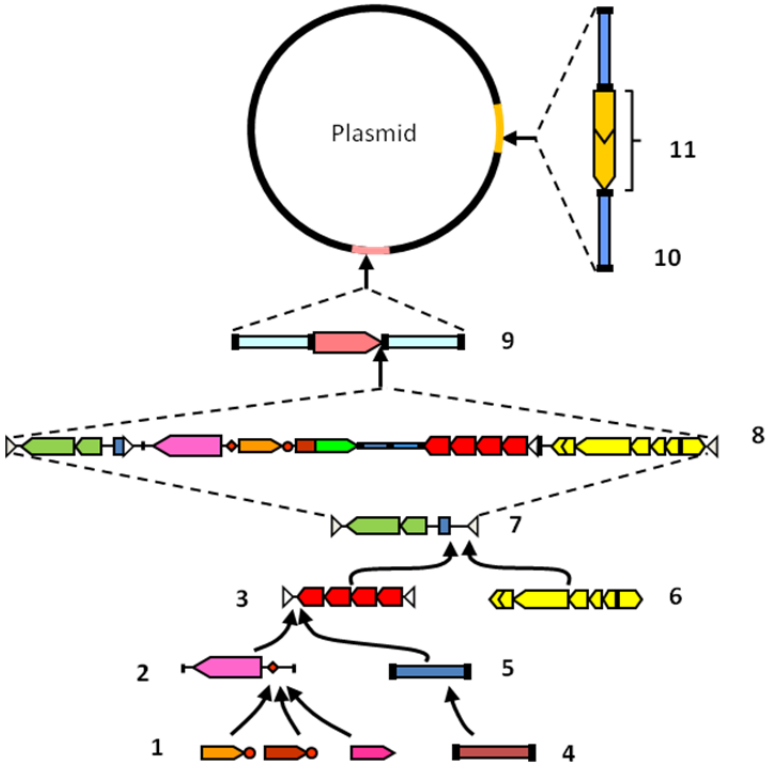


Figure 5: The modular structure of mobile elements. DNA elements that have fixed in populations as a consequence of human use of antimicrobial agents often have a modular structure, being assembled from diverse genetic events, each with a different evolutionary history. This schematic presents an overview of the modules that comprise an putative self-transmissible plasmid containing multiple resistance determinants. Gene cassettes encoding resistance genes (1) captured by Integron (2), which itself inserted into transposon (3). Subsequently, this transposon (3) was modified by insertion of two insertion elements (4, 5) and further inserted into another transposon backbone (6) that carried a resistance operon (7) generating another transposon (8) which in turn was itself inserted in another transposon (9). In turn this transposon is carried in a plasmid which also carries another transposon (10) conferring resistance to a distinct antibiotic (11). Adapted from (Gillings and Stokes, 2012).

With time and ongoing selection, the evolution trend is towards increasing complexity of multidrug resistance elements. This increase in complexity is driven by combinatorial exchanges between existing elements, recruitment of new elements (Walsh, 2013) and co-selection of genes which confer resistance to environmental compounds, pollutants and antibiotics (Baker-Austin et al., 2006). Isolated or grouped, these multidrug resistance elements follow a hierarchical organization and can be transferred preferentially among groups of phylogenetically related organisms that share a particular

environment (Baquero, 2004). In general, changes in a local population may contribute to the selection of advantages for linked organisms increasing the size of the gene pool which may drives antimicrobial resistance (Reanney, 1976).

Functional metagenomics reveals ready-to-use antimicrobials resistance genes

To analyze the microbial communities, and the effects of different injuries (as the presence of antibiotics) in their population dynamics, much progress has been made using 16S sequencing, direct bacterial culture, as well as shotgun metagenomic sequencing, together with PCR probing for specific genes, chemical profiling of microbial metabolites and detection of mobile elements (Bengtsson-Palme et al., 2014). However, sequence-based methods do not provide information on novel activities (as new antibiotic resistance genes) for which sequence information is not available.

Firstly suggested in 1985; the direct cloning of environmental DNA to classify uncultivable microorganisms (Pace et al., 1986) and the first successful function-driven screening of metagenomic libraries conducted in 1995 (Healy et al., 1995), are the parents of functional metagenomics approaches. Knowledge about the diversity and function of uncultured microorganisms containing resistance genes which are present in distinct environments and the connections among them is important to predict the emergence of resistance in bacterial populations. However, the mechanism behind the selection and spread of these elements remain to be clearly established. Large-scale projects have been used to advance the knowledge of organism and environmental microbiome and track the presence of those specific genes in such metagenomes (Handelsman and Tiedje, 2007). To this, meta-omics projects are the best option to investigate the behavior of uncultured organisms present in the environment and functional metagenome is the only method available which allows the discovery of novel antibiotic resistance (or any other functionally testable) genes and isolate mobile genetic elements that would not have been isolated in any other ways (Mullany, 2014; Torres-Cortés et al., 2011).

In countries where the antibiotic usage is not well regulated, the overuse of antibiotics has become a global health problem because the majority of chemicals compounds is partially absorbed or metabolized being discharged as draining waters from houses, hospitals, agriculture and farms. Associated with the capability to exchange DNA via lateral transfer of mobile genetic elements (e.g., plasmids or transposons), antimicrobial resistance genes from non-pathogenic and non-cultivated environmental microorganisms have the potential to be transferred, especially when use of antibiotics imposes heavy selective pressure. However, discovery of new AR mechanisms is limited because most studies focus on cultivable bacteria and known genes detected via PCR resulting in an incomplete understanding of the broader environmental resistome (Martiny et al., 2011).

Functional metagenomic screenings, originally proposed as a method to characterize the uncultivable fraction of soil microbiota (Handelsman et al., 1998) and the

functional diversity of microbes in many environments have been adapted to different purposes including the study of resistance genes (Allen et al., 2009; Berlemont et al., 2009; Torres-Cortés et al., 2011). This method does not require direct culture of isolated organisms but clone genomic libraries are constructed and then the fragmented DNA cloning into a vector screened for the desired function (Figure 6). This approach has the ability to reveal genes (novel or not) conferring resistance to a particular antibiotic, without requiring prior knowledge of the gene classes of interest. In contrast, most studies of antimicrobial resistance have used PCR to characterize specific AR genes, often known from clinically relevant bacterial isolates (Martiny et al., 2011).

In this regard, the functional metagenomic approach (Figure 6) has, considering its limitations, several advantages over a straight sequencing-based metagenomic analysis of a community because selection may be directly guided and allows screening of rare and novel genes in the community (Sommer et al., 2009).

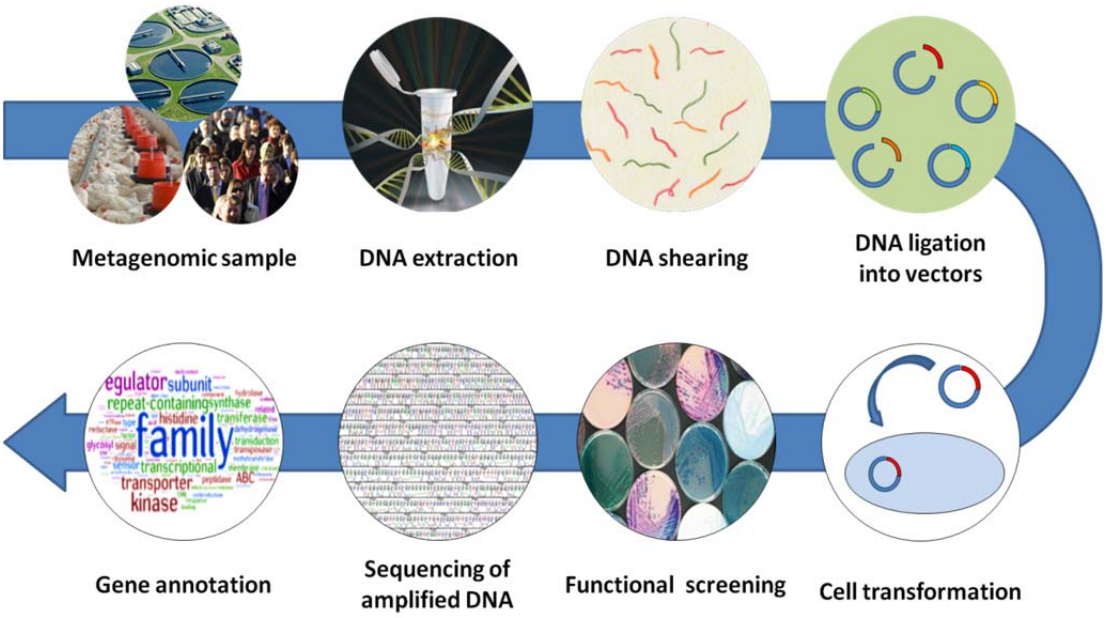


Figure 6: Workflow of the functional metagenomic analysis for characterizing antibiotic resistance genes from metagenomic samples. Metagenomic cloned cells harboring DNA fragments encoding antibiotic resistance genes are selected by subjecting the library of clones to specific antibiotics. Selected DNA fragments can then be sequenced to identify the specific resistance genes (Adapted from Sommer and Dantas, 2001).

As stated, functional metagenomic analysis is the only type of metagenomic screen that has the ability to isolate completely novel antibiotic resistance genes. However the disadvantage is that the resistance genes must be expressed and few species are currently in use as host of cloned genes. When such hosts are intrinsically resistant to a given antibiotic, functional screening is not possible. On the other hand, genes which may no confer resistance on their original host may, may render resistance, upon their over-expression driven by an heterologous promoter in the new host (Mullany, 2014; Schmieder and Edwards, 2012).

Stenotrophomonas maltophilia: an environmentally-born intrinsically resistant pathogen

Taxonomically, the genus *Stenotrophomonas* belongs to a subclass of proteobacteria firstly described as *Pseudomonas maltophilia* at 1961. Further it was reclassified as *Xanthomonas maltophilia* and finally classified as *Stenotrophomonas maltophilia* (Hugh and Ryschenkow, 1961; Palleroni and Bradbury, 1993; Swings et al., 1983). The exclusivity of nutrition requirements as suggests in its name, highlight the restriction of the nutritional range (from the Greek '*stenos*', meaning narrow, '*trophus*', meaning one who feeds and '*monas*', meaning unit) (Palleroni and Bradbury, 1993). Nevertheless, the finding of *S. maltophilia* in several different ecosystems supports that this narrow nutritional range does not imply a restriction of niche.

Indeed, different strains of *S. maltophilia* have been found in many environmental samples. This species is preferentially found in natural environments, being recovered from habitats like the rhizosphere of plants, animals and foods. In addition, *S. maltophilia* does not infect healthy people, but can cause infections in immunosuppressed patients, as those infected with HIV virus or under anticancer therapy. These infections include respiratory tract, soft tissues and bones, blood, eye, heart and brain infections among others, as cystic fibrosis patients and patients with previous therapy with broad-spectrum antibiotics (Adamek et al., 2011; Bader et al., 1999; Berg et al., 2001, 1999; García et al., 2008; Hilker et al., 2014; Lira et al., 2012; Pankuch et al., 1994; Zhang et al., 2001; Zhu et al., 2012). *S. maltophilia* infections are associated with high mortality in severely debilitated individuals demonstrating the need of strategies to improve the clinical treatment of infected patients especially in the intensive care units (ICUs) (Paez et al., 2008).

Apart of the clinical cases described above, this species exerts an extraordinary range of activities, including beneficial effects for plants, increasing their growth and protecting them from infections by microbial pathogens, its use for the biodegradation of pollutants or the production of molecules with economic value (Ryan et al., 2009). One important issue to analyze then is whether infective and natural (non-clinical) *S. maltophilia* isolates could constitute two different evolutionary branches in this species or rather if any strain can infect a compromised person. This is particularly relevant in order to evaluate the risks for human health associated to the use of beneficial *S. maltophilia* strains in agriculture purposes. In addition since this species is recognized as a model of intrinsically-resistant microorganism (Lira et al., 2012), studying the distribution of resistance determinants among different isolates of this bacterial species may help in the understanding of evolutionary forces driving the development of antibiotic resistance in bacterial recognized as pathogenic.

In order to understand the heterogeneity among *S. maltophilia* isolates, many authors have performed analyses as pulse-field gel electrophoresis (PFGE), random

amplified polymorphic DNA (RAPD) and intergenic consensus sequence-PCR (ERIC-PCR). Results from these studies suggest that clinical isolates of this species have more chances to adapt to the environmental conditions that surround them (Brooke, 2012). Concerning antibiotic resistance, the discussion about who came first, as the chicken or the egg, is relevant in this context. In other words, are clinical *S. maltophilia* isolates resistant because of the selective pressure exerted at clinical settings or rather the intrinsic resistance shared by the different *S. maltophilia* isolates constitute a major element for its success for causing infections at clinical settings?

The complete genome sequence of two clinical strains, *S. maltophilia* K279a (Crossman et al., 2008) and D457 (Lira et al., 2012) and two environmental strains, *S. maltophilia* R551-3 (Lucas et al., 2008) and JV3 (Lucas et al., 2011), indicate that several genes responsible for intrinsic antibiotic resistance are shared by all isolates independently of their habitats indicating a propagation and evolution of these elements before the intense use of antibiotics for humans, plants and animals treatment (Adamek et al., 2011; Alonso and Martínez, 2000; Crossman et al., 2008). Nevertheless, despite the presence of this set of genes in all genomes of strains from different origins, a clear genomic phylogenetic study comparing environmental and clinical isolates is still absent. According to this, the knowledge about the relationship between the isolates of *S. maltophilia* comparing the genomic repertoire of each one will enlighten relevant questions about their evolution, conserved genes and the range of occurrence based on the gene composition lost or acquired through Horizontal Gene Transfer (HGT).

Due to the genetic diversity and adaptation to different environments observed among *Stenotrophomonas maltophilia* isolates, the approach of sequencing multiple strains from a single species whose has been lead for studying the gene repertoire of given species (Adamek et al., 2011; Bader et al., 1999; Berg et al., 2001, 1999; García et al., 2008; Hilker et al., 2014; Lira et al., 2012; Pankuch et al., 1994; Zhang et al., 2001; Zhu et al., 2012), might serve to increase our knowledge of the populations dynamics.

Because of this, in the present work, we focused at the sequencing of one clinical isolate used as model at our laboratory and other 20 isolates of *S. maltophilia*, 10 of them obtained from clinical environment and 10 from environmental samples. We analyzed those genomes together with three further complete genomes previously published (Table 2). The data generated by this study will be useful to create a profile of all strains and understand the most important mechanisms of resistance and virulence-related genes of closely strains bringing valuable data on the ecological dynamics and adaptability of *S. maltophilia* in different contexts.

Predicting the antibiotic resistance

The unique, and sensate way to minimize the consequences of the widespread use of antibiotics is the prevention. To this, different prediction methods try to anticipate the bacterial capability to confer resistance to antimicrobial compounds. These methods are

based in the study of the capability of bacteria to acquire resistance through mutations as well as for establishing the bottlenecks that modulate the transfer of resistance genes among organisms from different origins (clinical, animal, environmental).

The constant donations and receptions of modular DNA fragments present in GMEs exert an important role at the transmission of mobile antimicrobial resistance elements as they may occur many times. Taking part of these modules, genes which encode and regulate the expression of operative elements can mediate the required antibiotic resistance and the transfer of these elements in blocks enhancing the capability of HGT for conferring multidrug resistance (Baquero, 2004; Martínez et al., 2007).

Taking into account that the modularity of resistance elements is not a static event and the integration between transmissible elements from hosts to donors may be affected by the environment, in a natural habitat, the diversity of organisms, mobile elements and the composition of the microbiota where they are present may exert a positive effect enhancing the HGT enriching specific elements involved with the resistance (Alonso et al., 2001; Beaber et al., 2004). As consequence, the diversity and abundance of organisms in a given ecosystem can be measured and further used to predict the antimicrobial resistance based on the transmissible elements which compose the scenario (Allen et al., 2010).

Interestingly the interchange of elements is not so promiscuous as can be imagined. Incompatibilities as differences at the number of antimicrobial genes, the diversity of mobile elements as transposons and integrons, and the extent number of plasmids may restrict the transfer of the different modules. In this case, the fixation of stable modular architectures becomes difficult because functional interactions are limited. Then, the knowledge of the possible combinations, in time and space, permit to establish a network where the present elements are linked and may allow to predict the further steps at the resistance route (Martínez et al., 2007; Shapiro, 2005).

Understanding the mechanisms for the emergence, evolution and spread of antibiotic resistance requires a multi-level analysis approach, since the selection pressures operates at different levels, from genes to metagenomes. In agreement with this idea, the current thesis has been designed to cover the following objectives.

Objectives

"If one does not know to which port one is sailing, no wind is favorable."
Lucius Annaeus Seneca

2. Objectives

- Determine the fitness costs of plasmids containing quinolone resistance genes belonging to the pentapeptide repeats protein family.
- Determine the capability of quinolone resistance genes to evolve from low-level resistance to quinolones towards high-level resistance.
- Determine the population structure and composition of the *S. maltophilia* species. In particular analyzing whether or not infective and environmental strains constitute two different phylogenetic branches along the evolution of this species.
- Determine the effects of triclosan on the selection of antibiotic resistance genes and genetic mobile elements involved in antimicrobial resistance in sludge from waste water treatment plants.
- Determine the effects of triclosan in the taxonomic composition of the microbiota in natural environments.

Material and Methods

"Research is what I'm doing when I don't know what I'm doing."
Wernher von Braun

3. Material and Methods

Plasmids and bacterial strains

The 2,253 Kb plasmid vector pZE21-MCS1 (Lutz and Bujard, 1997), which contains a kanamycin resistance cassette, was used to clone all DNA fragments obtained by a functional metagenomic approach. The fragments were cloned at the *Hinc*II insertion point and the resulting plasmids used to transform the bacterial host Top10 (DH10B). Transformed bacteria were selected using kanamycin at 50µg/ml and ciprofloxacin to obtain quinolone resistance genes. A total of 13 different clones were obtained at the Laboratory of Drug Resistance and Community Dynamics of Technical University of Denmark (Table 3). The Minimum Inhibitory Concentrations (MIC) of all strains for ciprofloxacin were determined by susceptibility strip test in Mueller-Hinton (MH) agar plates.

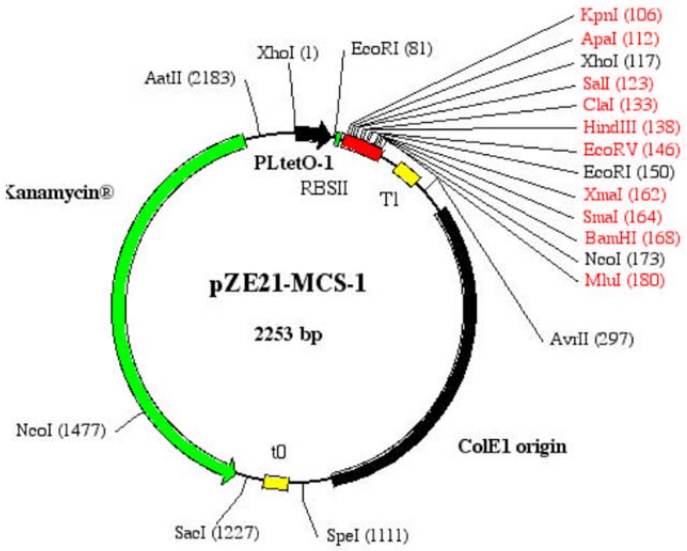


Figure 7: Representation of the plasmid vector pZE21-MSC1 with the kanamycin resistance gene in green.

Primers and amplification of inserted sequences

Primers for amplifying the cloned fragments by using the Polymerase Chain Reaction (PCR) were designed and purchased at Sigma-Aldrich (Saint Louis, Missouri). All amplifications made in this work were carried on using the DNA AmpliGel PLUS Master (Biotools, Spain) following the recommendations of the manufacturer. The PCR conditions used to amplify the insert cloned in the plasmid vector were as follows: 94°C for 5 minutes, (94°C for 1 minute, 60°C for 60 seconds, 72°C for 2 minutes) 25 cycles, 72°C for 7 minutes, 4°C hold. All PCR products and isolated plasmids were checked on a 2% agarose gel. The amplification and sequencing were performed using specific primers (Table 1).

Table 1: Primers used to amplify the insert fragments cloned into the vector pZE21-MSC1.

Primers	Sequence	TM	%GC	MW
pZE21-MSC1_Fw	5'-GATAGAGATACTGAGCACATCAG-3'	60.9°C	43.48%	7399
pZE21-MSC1_Rv	5'-TTTTATTTGATGCCTCTAGCACG-3'	63.9°C	39.13%	6999

Sequencing of quinolones resistance genes selected by functional metagenomic

To obtain the sequences of the genes conferring resistance to quinolones and track potential changes after experimental evolution (see below), the plasmids were extracted using the commercial system QIAprep Spin Miniprep Kit® (QUIAGEN) following the manufacturer's instructions and sequenced in both directions, using the same primers used for PCR (see above), by the facility of the Fundación Parque Científico de Madrid (Spain). The nucleotides sequences provided by the sequencing facility were analyzed using the software BioEdit (Hall, 1999) to assemble the sequences.

To know which gene was present in each clone, the sequences obtained from the sequencing were annotated with respect to their open reading frames (ORFs) (Figure 10), BLASTx was used to identify the genes present in the inserts and annotate the genes found. To identify the genes responsible for the resistance to ciprofloxacin, all consensus sequences were annotated using non-redundant (nr) databases.

Plasmid stability assay

The Top10 *Escherichia coli* clones containing the putative-quinolone resistance genes in the plasmid vector pZE21-MSC1 were pre-inoculated overnight in LB broth (Bertani, 1951) with kanamycin 50 µg/mL (37°C) and shaken at 250 rpm. Sequential overnight cultures were diluted 1/1000 into fresh LB broth without antibiotics and grown at 37°C at 250 rpm. Aliquots were removed every 24 hours and serially diluted (10^{-4}, 10^{-5}, 10^{-6} and 10^{-7}) and plated on LB agar plates in three replicates containing kanamycin 50µg/ml and three Petri plates without this antibiotic. The number of colony forming units (CFU) of each culture was calculate counting all colonies in both treatments (Figure 8A) and the percentage of plasmid-carrying cells between treated and non-treated plates were plotted every day (Figure 8B).

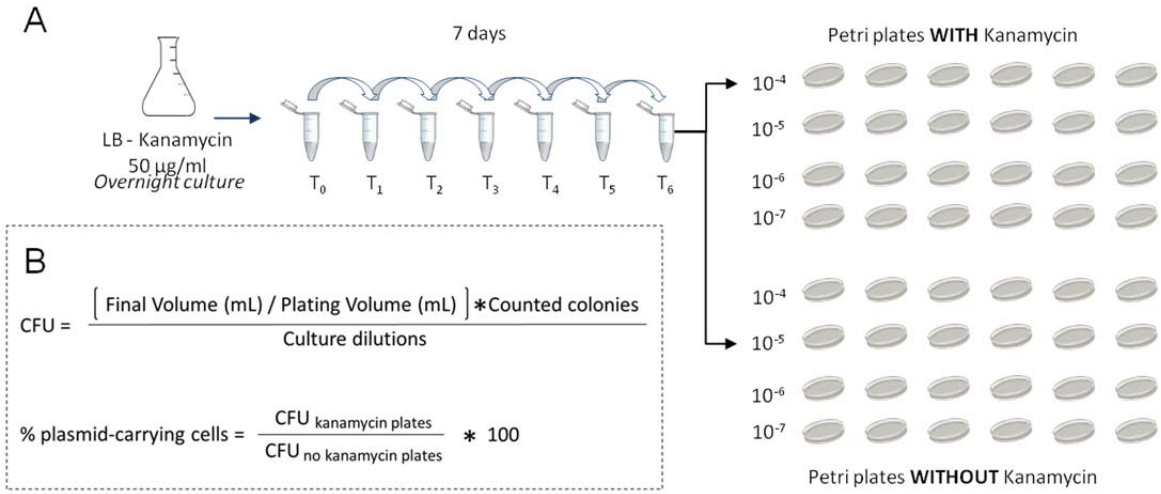


Figure 8: A) Representation of the experiment to verify the stability of the plasmid vectors pZM21-MSC1; B) Function to calculate the colony-forming units (CFU) from both treatments (plates with and without kanamycin) to check the plasmid stability; the percentage of remaining plasmid-carrying cells was determined by the second function inside the dotted box.

Determination of antibiotic susceptibility

For all strains the minimal inhibitory concentrations (MICs) were determined using Mueller-Hinton agar plates and strip tests following the manufacturer's instruction (Liofilchem®, Italy). In the case of bacteria carrying genes cloned in pZE21-MSC1, bacterial cultures were grown overnight (around 20h, 37°C) at 10ml of LB with 50µL/ml kanamycin. In the case of the *S. maltophilia* isolates, the overnight culture was performed without antibiotics. The OD_{600} of a dilution 1:10 (100µL culture: 900µL Mueller-Hinton Broth) of each culture was measured and cultures were diluted in NaCl 0,85% to an OD600 of 0,005. For all strains, 100µL of diluted cultures were spread in Mueller-Hinton agar plates and the strip test, containing the antibiotic of choice, was placed on the media surface as recommended by the manufacturer. After that the plates were incubated at 37°C overnight (±18 hours).

Plasmid evolution assays

To investigate the plasmid and bacterial host behavior in their adaptation to antibiotic increasing concentrations, the 13 strains were submitted to experimental evolution during 21 days using ciprofloxacin to select spontaneous mutants appearing after doubling the concentration of antibiotic in sequential passes (Figure 9). The initial concentration used for the evolution experiment was the already determined initial MIC for each of the clones. The plasmids extracted of the evolved clones were extracted using the same protocol mentioned before (QIAprep Spin Miniprep Kit®- QUIAGEN). To determine whether the changes in the susceptibility to quinolones were due to mutations in the plasmids or by contrary these mutations occurred in the chromosomes of the evolving clones, both evolved and original plasmids were introduced by transformation into *Escherichia coli* TG1 cells and the MICs of the transformants were determined for ciprofloxacin and norfloxacin using strip tests.

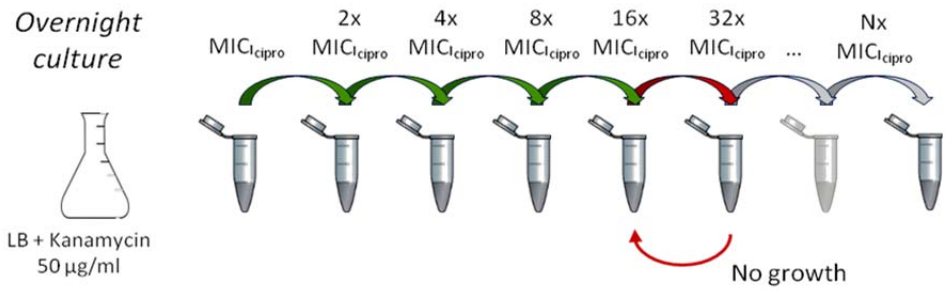


Figure 9: Representation of the evolution assay. All transformed cells were submitted to an overnight culture with a concentration of 50 µg/ml kanamycin increasing the concentration 2-fold each successive inoculation. When no growth was detected, the strain was submitted again to the previous concentration.

Sequencing, assembling and annotation of the whole genome of *S. maltophilia* D457

To access the genomic information of *S. maltophilia*, sequencing was performed using the 454 GS FLX system and single end and 3 Kb paired-ends reads. The total of bases obtained from both methodologies ranged in size from 64 to 147 Mb, respectively. The amount of reads for each sequencing method was of 370.000 and 700.000, for the single end reads and paired-end respectively.

Once received from the sequencing facility, the sequences were assembled *de novo* using the MIRA software (Chevreux et al., 1999) (http://chevreux.org/projects_mira.html) and revised using the Gap4 software (Staden et al., 2000). The genome of *S. maltophilia* K279a (Crossman et al., 2008) was used as reference to order the generated contigs. The gaps between the contigs as well as the repeated regions present in the genome were solved by PCR using primers from the end of the contigs and further Sanger sequencing of the amplicons. A total of 208 pairs of primers distributed along all the genome were designed exclusively to amplify the lacking fragments and clarify the position of repetitive zones. The coding sequences (CDS) of *S. maltophilia* D457 were initially annotated by using the RAST server (Aziz et al., 2008). Initial annotation was followed by a manual curation to confirm pseudogenes. The assembly inspection removed hypothetical coding genes smaller than 250 nucleotides without similarity with other *S. maltophilia* strains. Non-coding RNA genes were annotated by using several methods (Griffiths-Jones et al., 2005; Lowe and Eddy, 1997; Silva et al., 2006). Repeated sequences were analyzed to identify transposable elements and for mapping the inverted repeats at the end of the IS elements.

DNA extraction and genome sequencing of 20 new strains of *S. maltophilia*

A total of 20 isolates of *S. maltophilia*, 10 clinical and 10 environmental ones, were sequenced in this study (Table 2).

The isolates FL1 – FL10 were obtained from the Laboratory of Opportunistic Pathogens at National Spanish Biotechnology Centre – CSIC, Spain) collected from different sources (urine, sputum, respiratory secretion, blood and pus from wound) at the Hospital of Mostoles (Madrid, Spain) previously described (Alonso and Martinez, 2001). The isolates FL11 – FL20, sampled from environmental sites like sewage, rhizosphere, brackish water and eye-care solution, were provided by the Institute of Environmental Biotechnology at Graz University of Technology (Graz, Austria).

The total genomic DNA of each isolate was extracted using the GENOME DNA Kit (MP Biomedicals LLC, Illkrich, France) following the manufacturer instructions and whole-genome sequencing was performed using Illumina MiSeq technology (Illumina, San Diego, USA) using DNA libraries with insertion size between 700-800bp to generate paired-end

reads with 260-300bp length. The DNA sequencing was made at the facility of the Madrid Science Park (Madrid, Spain).

Table 2: Clinical and environmental isolates sequenced exclusively for this study. A total of 20 *S. maltophilia* strains were analyzed from clinical and environmental habitat. FL1 – FL10: clinical strains; FL11-FL20: environmental strains.

Code	Strains	Sample	Origin	Reference
FL1	E-729	Urine	Mostoles, 1993	(Alonso and Martínez 2001)
FL2	E-759	Sputum	Mostoles, 1993	(Alonso and Martínez 2001)
FL3	E-999	Respiratory secretion	Mostoles, 1993	(Alonso and Martínez 2001)
FL4	G51	Blood	Mostoles, 1994	(Alonso and Martínez 2001)
FL5	E-301	Urine	Mostoles, 1992	(Alonso and Martínez 2001)
FL6	D-388	Urine	Mostoles, 1991	(Alonso and Martínez 2001)
FL7	E-861	Sputum	Mostoles, 1994	(Alonso and Martínez 2001)
FL8	C-357	Urine	Mostoles, 1991	(Alonso and Martínez 2001)
FL9	E-539	Pus from a wound	Mostoles, 1993	(Alonso and Martínez 2001)
FL10	E-824	Blood	Mostoles, 1993	(Alonso and Martínez 2001)
FL11	N-S26	-	-	Provided by Gabrielle Berg
FL12	E-P13	Rhizosphere of rape	Rostock, 1995	(Minkwitz and Berg 2001)
FL13	E-A22	Sewage	Braunschweig, 1999	(Minkwitz and Berg 2001)
FL14	E-A1	Brackish water	Zingst, 1996	(Minkwitz and Berg 2001)
FL15	PS-5	Rhizosphere of oilseed rape	-	(Berg, Marten, and Ballin 1996)
FL16	E-A23	Eye-care solution	Munich, 1999	(Bader et al. 1999)
FL17	E-P20	Rhizosphere of potato	Braunschweig, 1998	(Minkwitz and Berg 2001)
FL18	E-P5	Rhizosphere of Brassica napus L.	Rostock, 1995	(Minkwitz and Berg 2001)
FL19	E-A21	Sewage	Braunschweig, 1999	(Minkwitz and Berg 2001)
FL20	E-A63	Sewage	Braunschweig, 2000	Provided by Gabrielle Berg

Quality control of the sequences and assembling

To check the quality score of the 20 sets of sequences, all reads were submitted to the software FastQC v.0.11.2, which allows an overview of the quality and the identification of contaminant sequences present after sequencing. The contaminant sequences were removed using the package AlienTrimmer v.0.4.0 software (Criscuolo and Brisse, 2013) with modifications for the *k-mers* (–k=12), the PHRED quality (–q=30) and length of the reads (-l =50) for the downstream analysis resulting in reads with minimal length of 50 bp with the PHRED quality over 30. After sequence trimming and filtering, 59.033.648 remaining reads (average of 2.951.682,4/strain) were assembled using MIRA v.4.20 software (Chevreux et al., 1999) in a 32Gb laptop. The contigs with less than 500bp were discharged of the further steps. Twenty sets of contigs with the average G+C% content of 66,10%, N50 = 59.858,3 and 97,33% of reads assembled (57.459.275).

The previously sequenced complete genome of *S. maltophilia* D457 was used as reference to order the contigs of the 20 isolates analyzed in this work. Three other complete genome sequences of *S. maltophilia* (K279, R551-3, JV3) available at NCBI were included in this study for comparative porpoises (http://www.ncbi.nlm.nih.gov/genome).

Comparative genomics

The interactive cross-platform BLAST Ring Image Generator (BRIG) (Alikhan et al., 2011) was used to map and visualize the contigs of the 20 sequenced isolates and the three other available complete genomes of *S. maltophilia* (K279, R551-3 and JV3). The genome of *S. maltophilia* D457 served as reference to map and visualize and compare all genome.

ORF detection, gene prediction and annotation

The prediction and annotation of the Open Read Frames (ORFs) found at the resulting contigs from each sequenced strain was carried using the software Prokka v1.11 (Seemann, 2014) with the following steps: 1) creation of a trustily customized database complied with all annotated genes from all available genomes of *S. maltophilia* (ftp.ncbi.nlm.nih.gov/genomes/Bacteria) which was used as the primary source of annotation using a expected value (*e-value*) threshold of 10^{-10} for BLASTp (Camacho et al., 2009); 2) BLASTp with approximately 16.000 proteins from UniProt (Apweiler et al., 2004); 3) BLAST+ with all proteins from finished genomes in RefSeq for specified genus; 4) hidden Markov model profile databases to perform a hmmscan using families profiles; 5) predicted genes were classified as "hypothetical protein" when no match could be found.

Genomic islands detection

Regions larger than 4000 bp with a minimum of four ORFs were classified as Genomic islands (GEI) and detected using the IslandViewer software (Dhillon et al., 2015).

Determination of the core genome and pangenome of four complete genomes of *S. maltophilia*

The core-genome (CG) and the accessory genome (AG) of the four *S. maltophilia* complete genomes were calculated using the GET_HOMOLOGUES package (Contreras-Moreira and Vinuesa, 2013). Clusters of homologous gene families were generated using the COGtriangles algorithm with coverage and identity set to 95% both. The soft core, shell and cloud genomes were calculated with the same software. In addition to the Perl script get_homologues.pl, a few auxiliary scripts helped with subsequent analysis parsing tables, generating figures and phylogenetic trees.

To estimate the preliminary core genome and pangenome sizes of *S. maltophilia* based on complete genomes, two clinical strains K279a (NC_010943.1) and D457 (NC_017671.1), and two environmental strains R551-3 (NC_011071.1) and JV3 (NC_015947.1), were used.

Further, the draft genomes of the 20 new strains were included at the analysis to estimate the core genome and the pangenome of the *S. maltophilia* species. The parameters to perform the analysis of the core genome for the 24 strains were established by an array of settings using different combinations of coverage and identity (coverage 85, 90, 95 and 98%; identity 70, 75, 80, 85, 90, 95 and 98%) to generate the clusters of homologues genes (Figure 23). Once the parameters were established, all genes were grouped into clusters with at least 90% of coverage and 95% of identity. The genes with <90% in coverage and <95% identity were assigned as accessory genome.

Functional annotation of *S. maltophilia* predicted genes

All predicted genes from accessory and core-genome were submitted to RAPsearch2 (Zhao et al., 2012) as a first step before establishing a functional annotation using the non-redundant NCBI database and further COG functional assignation using a workflow chart and complemented using the RAST server.

Origin, treatments, DNA extraction and sequencing of sludge samples

Two distinct samples of sludge (S1 and S2) collected from a Waste Water Treatment Plant (WWTP) in Turkey were submitted to two regimens of triclosan exposure using two microcosms named respectively Exp1 and Exp2. These microcosms were created to investigate the effects of triclosan when exposed at low and high concentrations over environmental samples to simulate it effect in natural conditions. S1 and S2 samples were cultivated under aerobic conditions. For S1, subsamples submitted to an increasing of triclosan exposure starting with 100 ppt, passing to 100 ppb, 500 ppb, 1 ppm until to reach the final concentration of 10ppm, were collected from Exp1 in intervals of seven days. This proceed was repeated during six weeks. For microcosm Exp2 a different approach was used; two subsamples of Exp2 were exposed to two distinct concentrations of triclosan (1 ppm and 10 ppm) without intermediate steps during seven days. For both regimens a control sample were collected at the beginning of the experiment, Control I and Control II to know the initial composition of these samples. This part of the study was performed in collaboration with Dr. Ülkü Yetiş from Middle East Technical University (Turkey).

The total DNA of each sample collected in this study was extracted using the PowerSoil® DNA Isolation Kit (MOBio Laboratories, Inc. - Carlsbad, CA USA) kit following the manufacturer's instructions. Total DNA sequencing was made in two separate stages at the facilities of the Madrid Science Park (Madrid, Spain). In the first stage, three sample from Exp1 (Control I, 100 ppb and 10 ppm) and two of Exp2 (Control II and 10 ppm) were sequenced using Illumina GAiix Technology (Illumina, San Diego, USA) with paired-end reads (2x100) with a insertion size ranging between 100 – 195 bp. For the second stage, the remaining samples of Exp1 (100 ppt, 500 ppb and 1 ppm) and Exp2 (1 ppm) were sequenced using Illumina MiSeq Technology generating paired-end reads with 300 bp in length (2x300) with the insertion size ranging between 225 – 275 bp. All sequences were analyzed using the Bioanalyzer Trace after Amplicon PCR Step.

Quality control of total metagenomic sequences

To check the quality score of the sequences generated from the nine samples, all reads were submitted to the software FastQC v.0.11.2, which allowed the identification of the remaining contaminant sequences. To filter low quality reads and to clean extremes with unrecognized nucleotides, all sequences were analyzed using the script IlluQC_PRLL.pl embedded at NGS QC Toolkit v2.2.3 package with the cut-off value for

PHRED quality score established to 30 resulting in high quality paired reads (Patel and Jain, 2012).

Identification of resistance genes and mobile elements into the metagenomes

Four customized databases were constructed separately to investigate the effects of triclosan in the composition and distribution of mobile elements present in each sample collected at different concentrations of triclosan. The databases were composed by sequences from plasmids genomes obtained from the National Center for Biotechnological Information (NCBI) (ftp.ncbi.nlm.nih.gov/genomes/Plasmids/), antimicrobial resistance genes collected from the Comprehensive Antibiotic Resistance Database (CARD) (http://arpcard.mcmaster.ca/blast/db/), resistance gene cassettes obtained from the Repository of Antibiotic resistance Cassettes (RAC) (http://rac.aihi.mq.edu.au/rac/) and integrons extracted from 'The Integron Database' INTEGRALL (http://integrall.bio.ua.pt/). All sequences used to construct the databases were downloaded at November 2013.

Before proceeding with the identification of the sequences, all bidirectional reads (bireads) were overlapped in the extremes with the intention to create a large sequence increasing the length of sequence to be used in the identification of the mobile elements. The Basic Local Alignment Tool (BLAST) was used to identify and annotate the bireads according their similarity to the database's sequences with an expectation value (e-value) of 10^{-10}. Passed this step, the composition was determined by the number of unique sequences which matched with at least one sequence, the distribution was determined by the presence or absence of one sequence in one determined sample, and the abundance of each kind of mobile element were determined dividing the number of bireads matched with the database sequences with the total number of bireads present in each sample.

Determination of the microbial composition of metagenomic samples

To access the information about the taxa composition and abundance of each taxa in the metagenomic samples extracted from each step of Exp1 and Exp2, a variable section of the 16S genes were sequenced. To this end, the V4 region was amplified from the total rRNA obtained from each microcosm sample and sequenced using Illumina Technology HiSeq with paired-end reads (2x300) generating reads with an average length of 225 and 255 bp.

All resultant reads originated from the sequencing were analyzed using the Mothur Software vs. 1.36.0 following the steps of the Standard Operating Procedure (SOP) optimized for Illumina sequences to identify the Operational Taxonomic Units (OTUs) using the Silva 16S rRNA database for bacteria references with 14,956 sequences (Kozich et al., 2013). The variation in composition and abundance of each sample were represented at Phylum, Order, Family and Genus levels.

Results

"*However beautiful the strategy, you should occasionally look at the results*."
Winston Churchill

4. Results

Capability of evolution of low-level resistance towards high-level resistance of plasmid-encoded *qnr* genes

The fact that antibiotic resistance genes are present in microorganisms that do not produce antibiotics has raised the question on whether there exist in nature pre-resistance genes existing in non-clinical antibiotic concentrations that, upon selective pressure in clinical ecosystems can confer adaptive phenotypic and genotypic responses to high-level resistance (Aminov, 2009). In order to test this possibility, a collection of thirteen clones containing DNA fragments, from different origins conferring low-level resistance to ciprofloxacin was selected by functional metagenomics from different samples in collaboration with Professor Morten Sommer. The libraries were constructed using the vector pZE21-MSC and cloned into *E. coli* Top10.

The initial MIC of this antibiotic for each strain was measured using strip tests (Table 1). The values demonstrated a variation of 8x between the lowest (0,250 µg/mL) and highest (2 µg/mL) MIC found. The clone Cip1 showed the higher MIC value (2 µg/mL). For the others clones MIC values varied between 0,25-1 µg/mL of ciprofloxacin.

All clones were sequenced and annotated to determine the content of the insert. The size, organization of the ORFs and annotations of each cloned fragment are summarized in **¡Error! No se encuentra el origen de la referencia.**. Some of the clones, presented the same insert (Cip1 and Cip5; Cip2 and Cip7; Cip3, Cip6 and Cip9) and in most cases, the plasmids harbored genes belonging to the pentapeptide repeats family, which comprises among others, proteins from the Qnr family, which confer low-level resistance to quinolones.

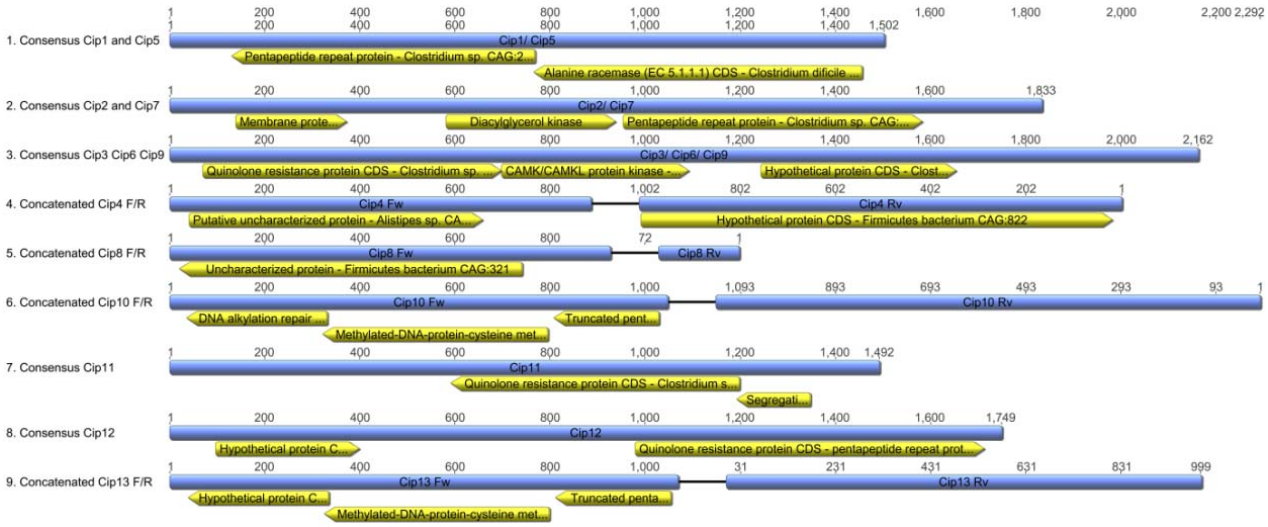


Figure 10: Representation of the inserts cloned demonstrating the organization of the ORFs and annotations of the genes present in each one.

Table 3: Minimum Inhibitory Concentrations (MICs) for ciprofloxacin measured for all samples at the beginning of the study with the selected fragments obtained from functional metagenomic samples selected with ciprofloxacin.

Clones/Fragments	Sample origin/code	Ciprofloxacin MIC (μg/ml)
Empty plasmid	*pZE21-MSC*	0,0625
Cip1	*HuGranja-cip-B1*	2,000
Cip2	*HuGranja -cip-C1*	0,250
Cip3	*KalveGranja-cip-C2*	0,250
Cip4	*KalveGranja -cip-E02*	0,250
Cip5	*HuGranja -cip-E1*	0,250
Cip6	*KalveGranja -cip-F2*	0,250
Cip7	*HuGranja -cip-A02*	0,250
Cip8	*Cow2Gup-cip-A02A03*	1,000
Cip9	*KalveGranja -cip-A04A05A07A09*	0,500
Cip10	*KalveGranja -cip-A08*	1,000
Cip11	*PigDKRI-cip-A11*	0,250
Cip12	*PigGranja-cip-A12B01B05*	1,000
Cip13	*PigGranja-cip-B02*	0,250

Plasmids stability

The maintenance and capability of evolution of a given gene may depend on the fitness cost it confers to its host. Because of this, a first goal of this work was to determine the stability of the plasmids containing the different inserts. Although some of the original clones harbored the same insert, we decided to conduct the experiments with all of them. The strains were cultured during seven days with daily sequential passages. The percentage of stable plasmids to each strain was obtained from the ratio of remaining CFU in LB agar plates with 50μg/ml of kanamycin and CFU from plates without this antibiotic. This ratio was calculated along seven days and is show in Figure 11.

Some clones demonstrated to be less stable than others indicating that fitness costs depend on the gene present in the plasmid. After seven days, the most stable plasmids were Cip1, Cip2, Cip5 maintaining a percentage of plasmid-carrying cells above 50%. These data demonstrate a decreasing in number of cell-carrying plasmids along the experiment for eight strains with values under 50% of plasmid-carrying remaining cells. Cip1, 2, 5, 6 and 7 maintained their values over 60% The Cip4 clone seemed to be less stable losing their plasmids at day five, before the end of the experiment.

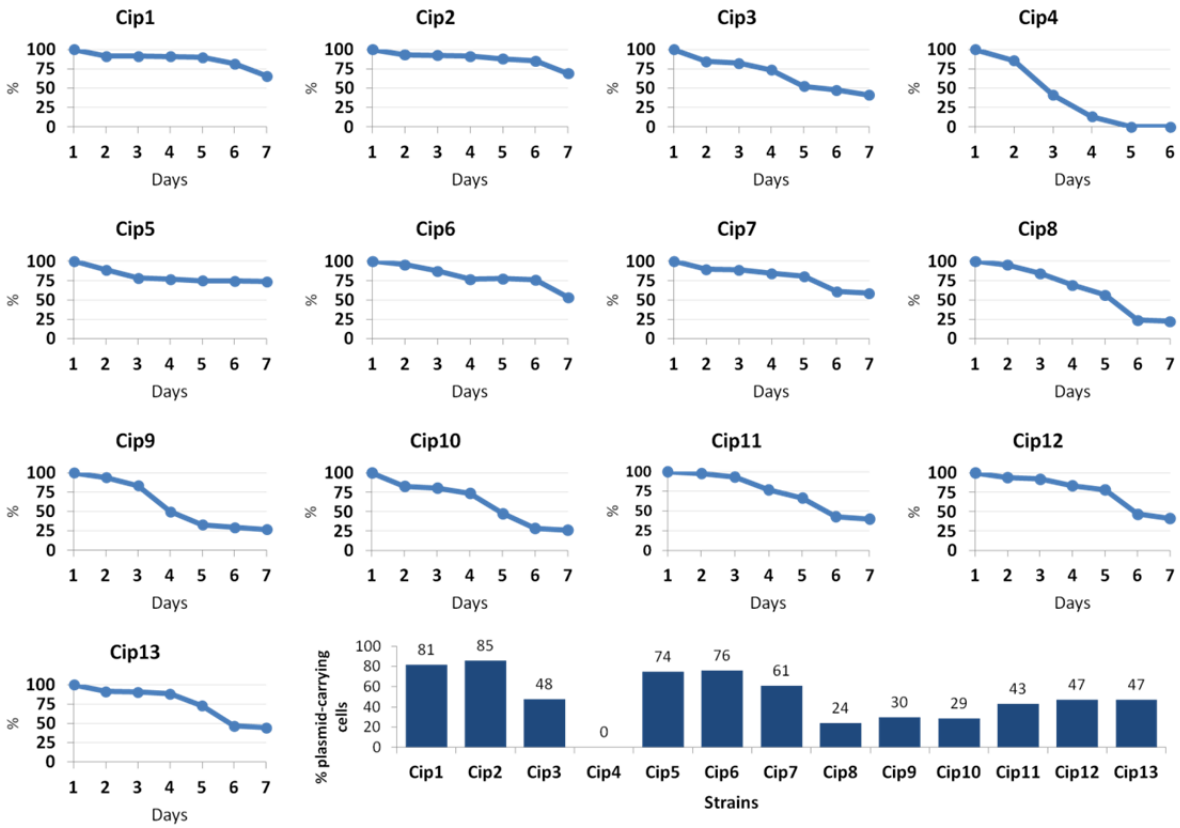


Figure 11: Plasmid stability along seven days calculated by the ratio between the Unit Formation Colony (UFC) present in Petri plates with 50µg/mL of kanamycin and plates without this antibiotic.

Experimental evolution towards high-level resistance: the same way do not reach the same destination

This experiment was conducted in two separate parts. In both experiments the 13 cloned strains (Cip1-Cip13) were evolved during 21 days submitting them to increasing concentrations of ciprofloxacin. The values for the MICs obtained previously for each strain by strip tests were the initial concentrations for the first cultivation of the evolution experiment. Cultures which did not grow were re-inoculated with samples grown in the previous concentration using the cells stored at -80°C obtained from the respective culture. The landscape of evolution after these 21 days was different for each of the strains (Figure 12). At the first experiment most strains grow despite the stepwise increasing concentration of ciprofloxacin. The Cip3 strain reached the highest resistance level growing at the concentration of 128 µg/mL of ciprofloxacin after 21 days representing an increasing of 520-fold compared with the initial MIC (0,250 µg/ml).

Curiously, the concentrations for Cip1 and Cip8 were reduced at the beginning of the experiment because these strains did not grow at the MIC determined previously, 2µg/mL and 1 µg/mL respectively. The selective concentration of Cip1 and Cip8 was reduced to 0,125 µg/mL of ciprofloxacin; however none of them crossed this threshold after 21 days of evolution. The maximal resistance level was reached in different times in

the different experiments and afterwards was maintained as constant for all strains varying from 12 (Cip7) to 21 days (Cip3, Cip5-6, Cip9-12).

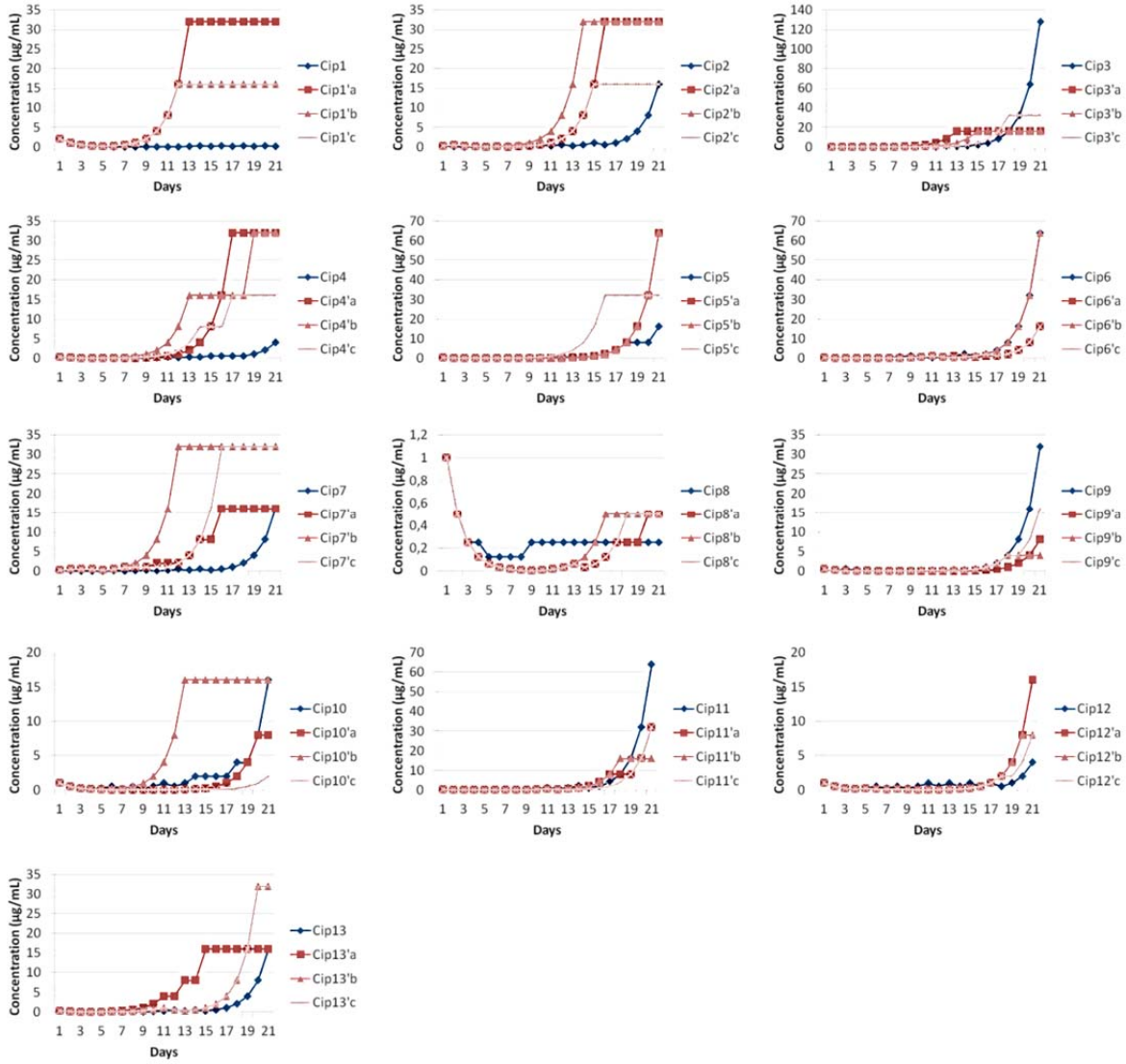


Figure 12: Plasmid evolutions demonstrating different behaviors and evolutionary story for each strain evolved independently. Blue lines represent the first experiment and red lines the second.

Re-transformation with the evolved plasmids do not preserve the previous resistance profile

Plasmids present in the evolved cells were extracted and cloned into *E. coli* TG1 to determine the MIC for ciprofloxacin. The values found with strip tests for this antibiotic demonstrated just a very minor difference between the MIC of the cells containing the original plasmids and the cells carrying the plasmids submitted to increasing concentrations of ciprofloxacin. Two strains, Cip7 and Cip13, presented slightly higher MICs for ciprofloxacin then the original plasmid-vector (Figure 13).

Because they demonstrated an increasing MIC for ciprofloxacin, MIC values were determined for four other quinolones to check if they presented the same behavior: nalidixic acid, norfloxacin, levofloxacin and ofloxacin. The evolved plasmids of both strains

presented slightly higher MIC values when compared with not-evolved plasmids. Original plasmid of Cip7 conferred a MIC of 16µg/ml while the value for the original plasmid was 12 µg/ml; Cip13 presented growth at 32 µg/ml which represents a concentration of 2-fold as compared with the strain carrying the original plasmid. With respect to the antibiotic norfloxacin, only the strain Cip7 presented a 2-fold MIC higher than the initial plasmid with no differences between the original plasmid and the plasmid of Cip13 submitted to evolution.

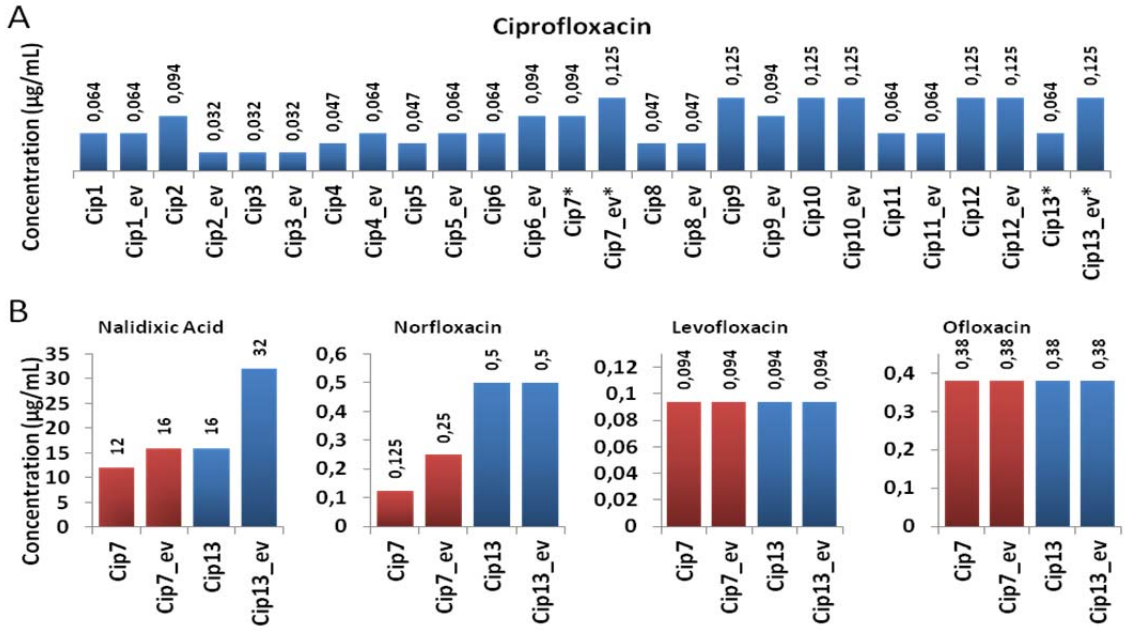


Figure 13: (A) Strip test results of the 13 strains re-transformed with the original and evolved plasmids; (B) Strip tests results of the selected strains Cip7 and Cip13 tested with Nalidixic Acid, Norfloxacin, Levofloxacin and Ofloxacin.

Based on the values obtained from the strip tests, both evolved and not-evolved plasmids were extracted from their hosts further sequenced using the plasmids mentioned at Table 1. The resulting sequences from the inserts were aligned to verify the presence of Single Nucleotides Polymorphisms (SNP) between the sequences of each strain. The alignment of the original sequences Cip7 and Cip13 with their respective evolved inserts presented silent mutations in their amino acids sequences. For Cip7, the nucleotide polymorphism A →G at position 18 of the gene did not make any change at the amino acid sequence when translated. Although, when analyzing the nucleotide sequences from the original and evolved plasmids we found three nucleotides variations. Translating the nucleotide sequence two of the three polymorphisms codified the same amino acid then the original sequence, been considered a compensatory mutation not exerting any influence at the protein sequence. The mutation present at the base 29 at the nucleotide sequence of the gene presented a change at the amino acid codified by the corresponding codon. This change was evaluated by the PROVEAN server (Choi and Chan, 2015) where the software detected a neutral mutation substituting the serine amino acid to a threonine residue.

These results indicate, even though some genes can evolve, their evolution renders just incremental changes in the MIC values. These results show that the high-level quinolone resistance achieved during the evolution experiment is due to mutation on chromosomal genes, likely on those encoding bacterial topoisomerases, and not in the genes present in the plasmids. It might be possible however that the presence of one gene or another promote such emergence, in which case the evolutionary landscape would be different for each of the clones (as shown in Figure 14, blue lines), been reproducible. To address this possibility, a second evolution experiment, in this case in triplicate was performed. As shown in Figure 14, evolution landscapes were not reproducible, indicating that the timing for the selection of the quinolone resistance chromosomal mutation does not depend on the gene present in the plasmid and is stochastic.

Figure 14: Single nucleotides polymorphisms found at the evolved inserts from Cip7 and Cip13. Cip7 (A) presented a polymorphism at the nucleotide 18 (A→G) of the ciprofloxacin gene and did not represented a representative mutation because coded the same amino acid as the original sequence. At Cip13 (B) mutations at the nucleotides 96 and 129 did not interfered at the amino acid codified by the codons which suffered changes. The polymorphism detected at the nucleotide 29 promoted a change at the amino acid sequence when compared with the original sequence not evolved. The substitution of the serine to a threonine amino acid was considered Neutral by PROVEAN software (B.1) (Choi and Chan, 2015).

Evolutionary trajectories of plasmid-encoded qnr genes present in human pathogens

The results previously shown suggest that plasmid-encoded Qnr determinants, present low chances to evolve towards high-level resistance. However, different alleles of *qnr* genes have been described in plasmids present in human pathogens. Two possibilities can explain this situation: *qnr* genes have evolved after their acquisition and hence each of the families has a monophyletic origin or, alternatively, there have been multiple acquisition events of *qnr* genes, in which case a polyphyletic origin, with each *qnr* allele being independently acquired, can be inferred.

To address the ways of quinolone resistance genes (*qnr*) evolved and how they are present into plasmids, two reference sequences, *qnrA* (gi: 87043704) and *qnrB* (gi: 149395207), of the referred genes present in plasmids were analyzed with respect to the plasmid genomic structure, as well as the elements surrounding each of the alleles for each

of both *qnr* families. To investigate the evolutionary steps, the coding sequences (CDS) of 32 plasmids (*qnrA* - 11 plasmids; *qnrB* - 21 plasmids) were analyzed (Figure 15A, B). The plasmids containing the different *qnrA* alleles were distributed between seven bacterial species and presented two allelic variants: *qnrA1* represented by seven plasmids and *qnrA6* present in three of them. The plasmids carrying *qnrB* genes, distributed at seven bacterial species, contained four variants for this gene, *qnrB1* was present in eight plasmids, *qnrB2* was shared among four plasmids, *qnr6* in two and *qnrB19* was located in seven plasmids. (Figure 15B).

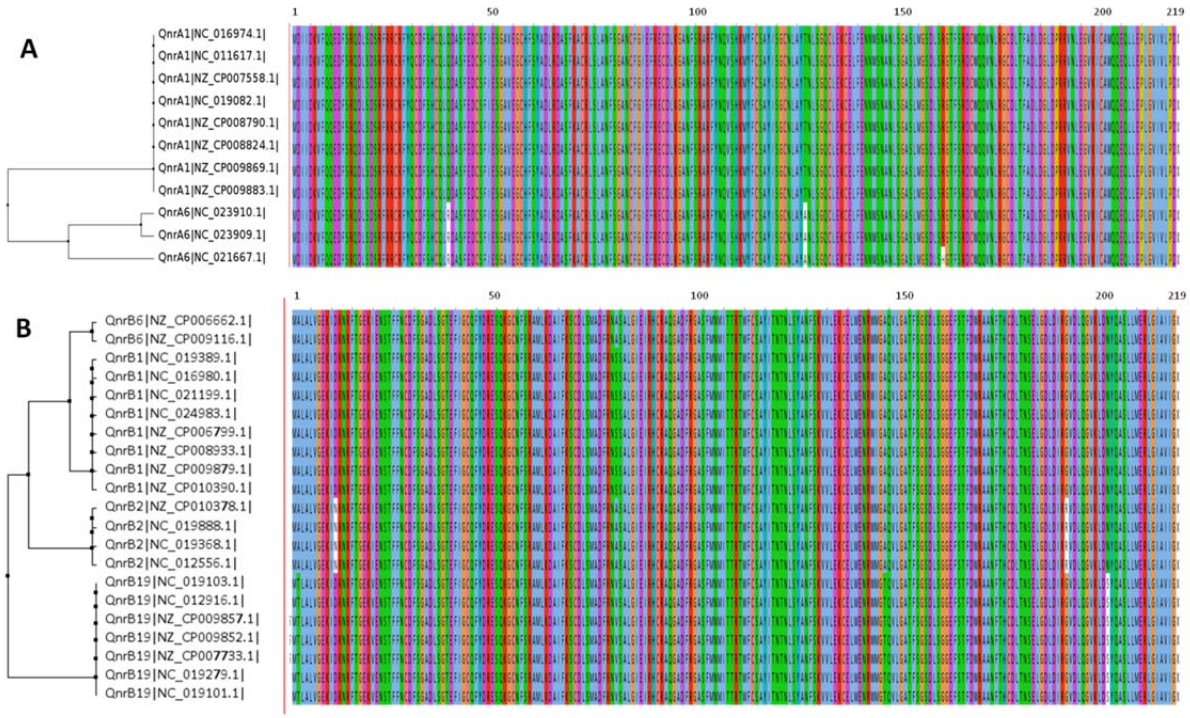


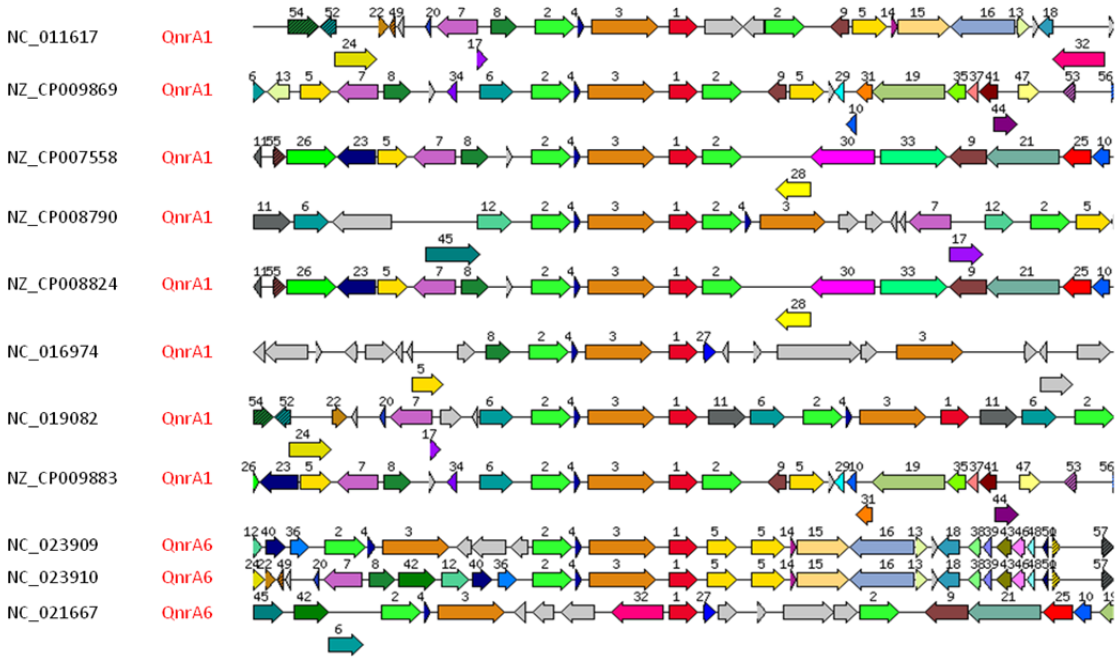
Figure 15: Alignment of the (CDS) amino acid sequences of the *qnr* genes: A - alignment of the QnrA products indicating the presence of two major groups containing the QnrA1 and QnrA6 variants; B - alignment of the QnrB products representing the four groups containing the variants QnrB1, QnrB2, QnrB6 and QnrB19.

QnrA1 encoding alleles usually presented a similar surrounding genomic structure, although the plasmids NC_011617, NC_016974 and, NC_019082 presented a diverse structure containing more genes inserted between the patterns mentioned before (Figure 16). Despite their similarity with the gene *qnrA1,* the *qnrA6* genes from plasmids NC_023909 and NC_023910 presented a completely different surround genetic structure. Inasmuch, the *qnrA6* allele presented a completely different surrounding structure in the plasmid NC_021667 (Figure 16).

QnrB plasmid-encoded variants presents a more diverse organization around them. The *qnrB1* allele presented a pattern composed by transposases surrounding the *qnrB1* gene. Plasmids carrying the *qnrB2* and *qnrB6* alleles presented similar organization concerning the position and composition around the genes. *qnrB19* allele was present in four plasmids containing only the *qnrB19* gene (NC_012916, NC_019101, NC_019103 and NC_019279) and other three with the genes inserted between two transposases

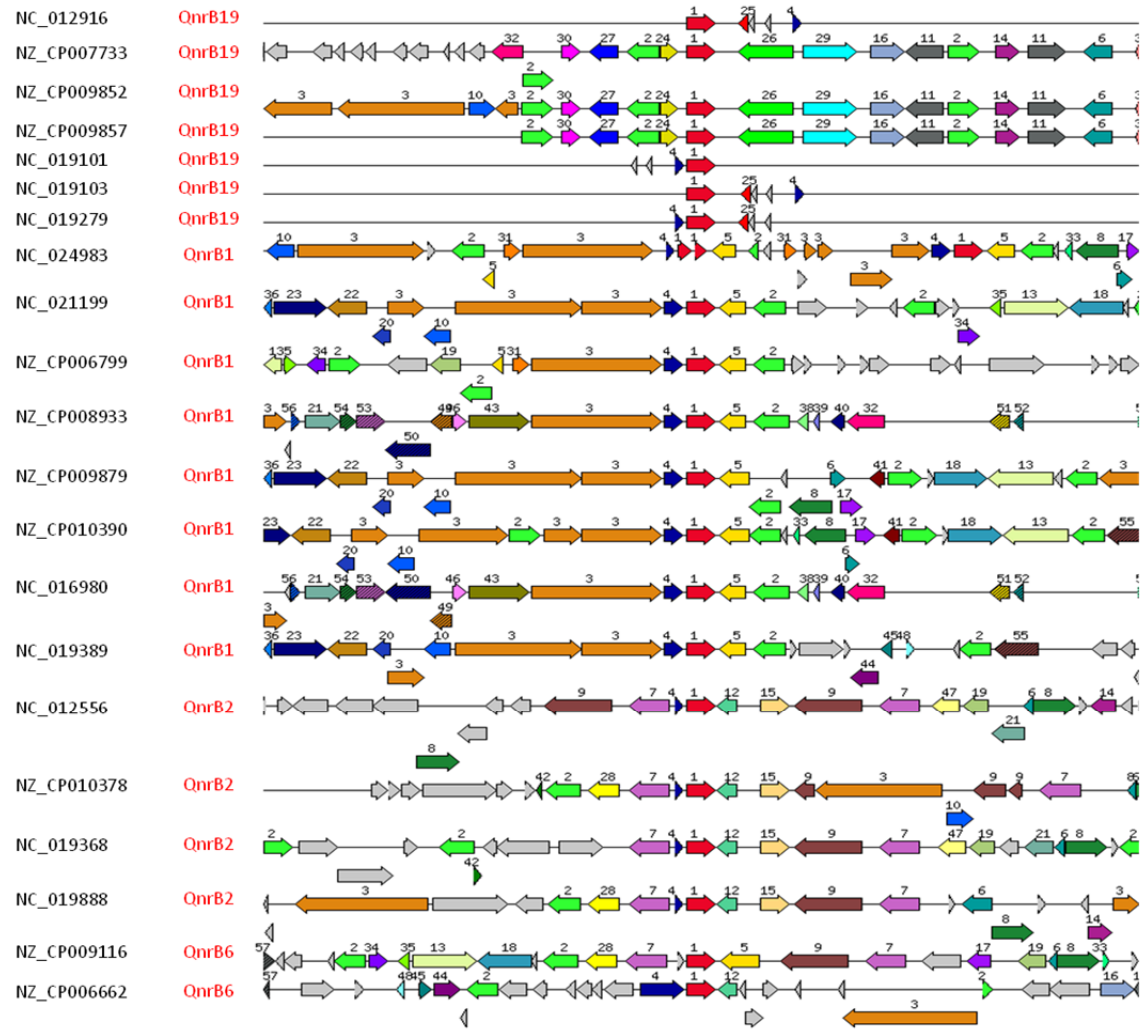
(NC_CP007733, NZ_CP009852 and NZ_CP009857) and presented a similar synteny in the surrounding region. (Figure 17).

Altogether these results indicate that the presence of different *qnr* alleles from both the *qnrA* and *qnrB* families is likely the result of different acquisition events and not of the selection, during treatment of infections with quinolones, of *qnr* variants conferring higher level of resistance. Altogether, these findings suggest that the different *qnrA* and *qnrB* genes alleles have been independently acquired, each allele flanked by different structures, supporting these alleles are not the result of evolution under selective pressure after their acquisition by human pathogens (Figure 16).



Nº	Function	Nº	Function	Nº	Function
1	Qnr	21	TniA putative transposase	41	Mercuric transport protein, MerT
2	Dihydropteroate synthase (EC 2.5.1.15)	22	hypothetical protein	42	Beta-lactamase (EC 3.5.2.6)
3	Mobile element protein	23	Lactose permease	43	Hypothetical protein
4	Hypothetical protein	24	Hypothetical protein	44	Mercuric resistance operon regulatory protein
5	Mobile element protein	25	Tn21	45	Chloramphenicol resistance protein
6	Beta-lactamase (EC 3.5.2.6)	26	Putative RecF protein	46	Hypothetical protein
7	Hypothetical protein	27	hypothetical protein	47	Mobile element protein
8	6'-N-acetyltransferase	28	Mobile element protein	48	Hypothetical protein
9	Mercuric ion reductase (EC 1.16.1.1)	29	Tn21	49	Hypothetical protein
10	Mercuric transport protein, MerE	30	Mobile element protein	50	Hypothetical protein
11	Beta-lactamase (EC 3.5.2.6)	31	Mercuric resistance operon coregulator	51	Hypothetical protein
12	Chloramphenicol acetyltransferase	32	Mobile element protein	52	Hypothetical protein
13	Mobile element protein	33	Mobile element protein	53	Hypothetical protein
14	TniB NTP-binding protein	34	Hypothetical protein	54	Replication initiation protein
15	Hypothetical protein	35	Mercuric transport protein, MerC	55	Hypothetical protein
16	TniB NTP-binding protein	36	Dihydrofolate reductase	56	Hypothetical protein
17	Spectinomycin 9-O-adenylyltransferase	37	Periplasmic mercury(+2) binding protein	57	ATPase
18	Probable membrane protein STY4561	38	Hypothetical protein		
19	Mercuric ion reductase (EC 1.16.1.1)	39	Hypothetical protein		
20	BsuBI-PstI family restriction endonuclease	40	rifampin ADP-ribosyltransferase		

Figure 16: Representation of the distribution and organization of the *qnr*A genes qnrA1 and qnrA6 present in the 11 plasmids analyzed. Each number indicates an arrow classified and colored by function. Red arrow (number 1) represents the *qnr*A gene.



Nº	Function	Nº	Function	Nº	Function	Size
1	Qnr	21	Beta-lactamase	41	TniB NTP-binding protein	354
2	Mobile element protein	22	Permease (DMT)	42	TniB NTP-binding protein	183
3	Mobile element protein	23	Tetracycline efflux protein TetA	43	Heat shock protein 60 family chaperone GroEL	1377
4	Psp operon transcriptional activator	24	Aerotaxis sensor receptor protein	44	FIG01046430: hypothetical protein	612
5	Oxidoreductase	25	hypothetical protein	45	putative cytoplasmic protein	282
6	Streptomycin 3"-O-adenylyltransferase	26	Mobile element protein	46	Heat shock protein 60 family co-chaperone GroES	291
7	Dihydropteroate synthase (EC 2.5.1.15)	27	Chloramphenicol acetyltransferase (EC 2.3.1.28)	47	Chloramphenicol acetyltransferase (EC 2.3.1.28)	633
8	hypothetical protein	28	hypothetical protein	48	hypothetical protein	183
9	Mobile element protein	29	Aerotaxis sensor receptor protein	49	hypothetical protein	504
10	DNA integration/recombination/invertion protein	30	YbjA protein	50	Cytochrome c-type biogenesis protein DsbD	1032
11	Beta-lactamase (EC 3.5.2.6)	31	FIG00638667: hypothetical protein	51	hypothetical protein	480
12	FIG00613169: hypothetical protein	32	hypothetical protein	52	hypothetical protein	207
13	DNA-cytosine methyltransferase (EC 2.1.1.37)	33	BsuBI-PstI family restriction endonuclease	53	Phosphoribosylanthranilate isomerase	639
14	Mobile element protein	34	hypothetical protein	54	bleomycin resistance protein	366
15	Peptide transport periplasmic protein sapA	35	Mobile element protein	55	replication initiation protein	1014
16	Glycerol-3-phosphate regulon repressor	36	tetracycline repressor protein	56	conserved hypothetical protein	195
17	Dihydrofolate reductase (EC 1.5.1.3)	37	hypothetical protein	57	Mobile element protein	252
18	FIG01046921: hypothetical protein	38	FIG00641267: hypothetical protein	58	RepFIB replication protein A	1128
19	6'-N-acetyltransferase	39	hypothetical protein			
20	Amidases related to nicotinamidase	40	hypothetical protein			

Figure 17: Representation of the distribution and organization of the qnrB genes qnrB1, qnrB2, qnrB6 and qnrB19 present in the 21 plasmids analyzed. Each number indicates an arrow classified and colored by function. Red arrow (number 1) represents the *qnr*B gene.

The whole genome sequence of *S. maltophilia* D457 a clinic model strain

At the beginning of this Thesis only three *S. maltophilia* strains have been fully sequenced and had their genomes closed, two environmental and one clinical isolates. We then decided to sequence another non-gapped reference genome from the model clinical isolate commonly used in our laboratory *S. maltophilia* D457.

The sequences received from the sequencing facility were used to generate the contigs and the first draft genome using other two other complete genomes as scaffold to

order the contigs. In a further step, in order to join the contigs and fill the gaps between the contigs we decided to amplify these regions, based on the information of the sequence scaffold. A total of 208 pairs of primers (Figure 18) were designed to amplify the missing fragments and clarify the position of repetitive zones present at the genome. After amplification, all amplicons were sequenced using Sanger technology and the information was used to close the complete genome of *S. maltophilia* D457.

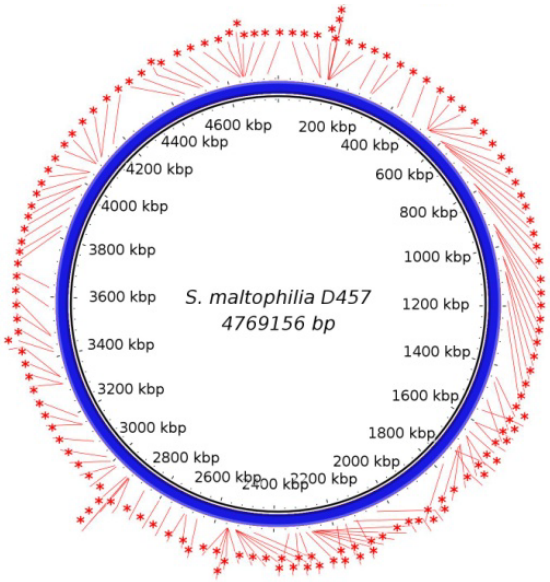


Figure 18: Representation of the 208 pair of primers used to close the genome of *S. maltophilia* D457. Each dot represents a primer designed to perform PCR reactions and sequence the products in order to fill the gaps between contigs.

The chromosome of *S. maltophilia* D457 comprises approximately 4,7Mb, (G+C content of 66.8%), a value in line with the average length of the genomes described for the other sequenced *S. maltophilia* strains. Its 4,209 genes are distributed in 4101 coding sequences (CDSs) and 108 non-coding RNA genes (13 rRNA, 71 tRNA, and 24 other RNAs). Six types of transposable elements were identified in this strain and only two of them (ISSmaD4 and Tn5044) were detected in other sequenced *S. maltophilia* genomes. Among the predicted coding sequences of D457, more than 200 genes were not shared with the other strains of *S. maltophilia* with completely sequenced genomes available.

Notably, most of the genes unique for *S. maltophilia* D457 encoded hypothetical proteins and transposases, which suggests that *S. maltophilia* has a large potential for gene mobility. The complete genome did not show chromosomal rearrangements when compared to those published strains (Figure 19) (Lira et al. 2012). In size, D457 and K279a had shown similar lengths with 4.769.156bp and 4.851.126bp, respectively; the environmental strains R551-3 and JV3 have slightly lower genomes with 4.573.969bp and 4.544.477bp, respectively.

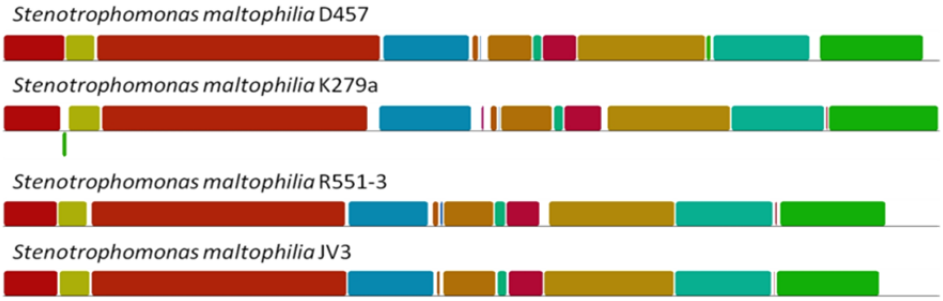


Figure 19: Multiple sequence alignment generated using the MAUVE software with the four complete genomes of *Stenotrophomonas maltophilia*. Colored blocks show conserved sequences at the genomes demonstrating the high similarity between the sequenced S. *maltophilia* D457 and the three available complete genomes.

The genomic islands of *S. maltophilia* D457 harbor pathogenicity determinants

Genomic Islands (GEIs) are regularly found in different genomes and frequently encode proteins responsible for the transfer of genetic material between organisms and the integration of these elements into GEIs, which is allowed often by the presence of flanking direct repeats. In addition GEIs usually present genes encoding recombinases that likely enable their integration into the host chromosome (Waldor, 2010). The analysis of *S. maltophilia* D457 showed the presence of some genomic regions distributed along its genome, which differs from the other sequenced strains. These fragments were recognized as GEIs. Because of the specific composition and segregation of functional modules by recombination, the GEIs of *S. maltophilia* D457 have regions that cannot be assigned to any other strain unambiguously, likely because of their transferable origin.

Observing in deep its complete genome, *S. maltophilia* D457 harbors 23 genomic islands (GEIs) with 277 genes present, among which, 106 are hypothetical proteins Table 4). All genomic islands in the D457 strain cover a total of 337.423 bp corresponding to 7% of its genome. with the largest one, GEI-5, containing 43.365 bp in length and 37 genes distributed along this genome section. When mapping on the *S. maltophilia* D457 genome, the three complete genomes (K279a, R551-3 and JV3) and the genomic islands from D457, it was noticed that not all genes present at the GEIs are exclusive from D457 with a few genes shared with other strains but not with all of them (Figure 20).

The characterization of all GEIs indicated the presence of genes involved in the metabolism of resistance to heavy metals (*cadA*, *copA, copB, copC, copD, copF, copG, cusA*, *czcA*, *czcD*, *merA*, *merP*, *merT*, *merR, smmJ*, *ssmK*, *smmP2*, *ssmQ2*), different genes associated with secretion systems (*tadA*, *tadB*, *tadC, virD*), conjugative transfer proteins (*tebC*, *tebD*, *trbB*, *trbC*, *trbD*, *trbE*, *trbJ*, *trbL*) as well transposases, transferases, kinases, proteases and integrases. The gene *sugE*, which confers resistance to a subset of quaternary ammonium compounds (QAC), was found at the genomic island GEI-8 (Appendix 3). The presence of heavy metal resistance genes in the genome could play a role in the ability to colonize environments where these compounds are present (Deng et al. 2014; Huang, Lu, and Chen 2013). Secretion systems are also important in the interaction of bacteria with different hosts, to permit the transport of effectors proteins through both

bacterial membranes. According to this, we detected the presence of type-1 (T1SS), type-2 (T2SS), type-4 (T4SS) and type-5 (T5SS) secretion systems in *S. maltophilia* D457 present at these islands.

Table 4: List of the 23 Genomic Islands (GEI) found in the genome of *S. maltophilia* D457

LABEL	START	END	SIZE
GEI-1	54822	63101	8279
GEI-2	162624	201509	38885
GEI-3	293704	299236	5532
GEI-4	317692	324915	7223
GEI-5	976021	1019386	43365
GEI-6	1027006	1031822	4816
GEI-7	1031382	1042897	11515
GEI-8	1055153	1074404	19251
GEI-9	1549512	1558514	9002
GEI-10	1648522	1669064	20542
GEI-11	1670959	1676372	5413
GEI-12	1693621	1698278	4657
GEI-13	1710319	1739152	28833
GEI-14	1761475	1774086	12611
GEI-15	1779635	1790559	10924
GEI-16	1955157	1966250	11093
GEI-17	2060200	2069193	8993
GEI-18	2424072	2437217	13145
GEI-19	2462648	2472668	10020
GEI-20	2509919	2517701	7782
GEI-21	3665701	3671533	5832
GEI-22	4197405	4233316	35911
GEI-23	4430040	4436910	6870

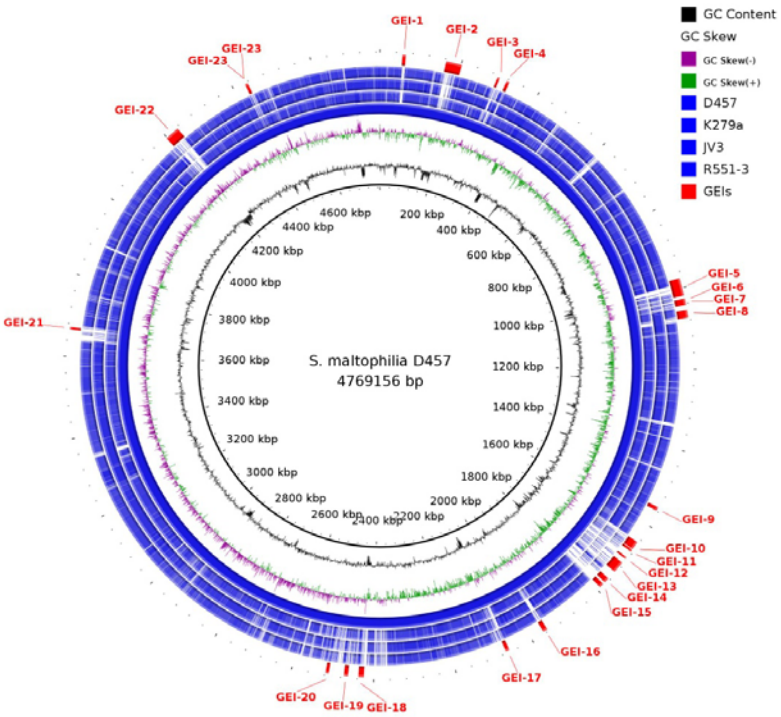
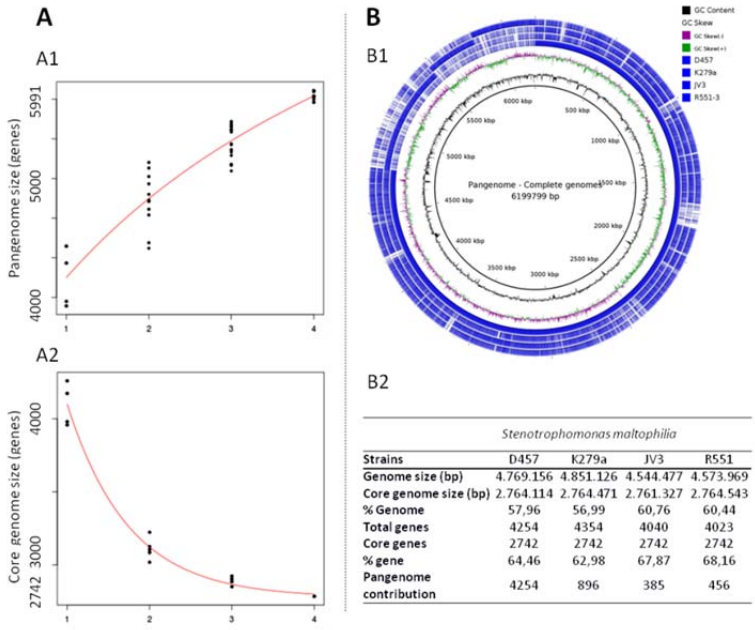


Figure 20: Circular representation of *S. maltophilia* D457 used as scaffold to align the three other complete genomes available using the BRIG software (Alikhan et al., 2011). From the inner to the most extern ring: 1 - GC% content; 2 - The clinical strain *S. maltophilia* K279a; 3 - environmental strain R551-3; 4 - environmental strain JV3; 5 – position of the 23 genomic islands present at the genome of *S. maltophilia* D457 present in regions that do not match with the other genomes of *S. maltophilia*.

Core and pan-genome of the complete genomes of *S. maltophilia*

The core genome consists in the set of genes shared by all individuals from a bacterial species. To investigate the differences at the genomic composition of the different strains, the whole genome alignment of the four strains was computed and the total lengths of the core genome and the pangenome were calculated. In total the core genome presented a coverage length of ≈2.764.363 bp and the pangenome covered a total of 6.199.799 bp. In order to define in more detail the genes shared by the two clinical (D457 and K279a) and the two environmental strains (R551-3 and JV3), the core genome was estimated based on a similarity of 95% and at least 98% of coverage level. In total, these four strains summed 5991 clusters of orthologous genes, of which 2742 were shared for all strains and constituted their core genome (Figure 21A).

The group of genes shared by just two strains or exclusively present in only one isolate has been defined as "cloud genome" (Kaas et al., 2012). In this group we included 3243 genes. This amount of exclusive genes not shared by all strains indicates that *S. maltophilia* harbors several strain specific genes. These genes could be responsible for the colonization and/or adaptation to different niches or hosts. For each genome, the core genome presented a length average of ≈2.764.363bp, covering in average 59,20% of the total genome. To determine the gene repertoire of all complete genomes, the number of genes added by each genomic sequence was estimated accumulating a total of 5991 genes that compose the pangenome for those strains. The contribution of each strain using the strain D457 as a scaffold (4.769.156 bp – 4254 CDS) was 896 when the second genome was added (K279a), decreasing to 385 genes with the third genome added (JV3) and finishing with 456 genes added by the strain R551-3 (Figure 21B). This finding indicates that the pangenome of *S. maltophilia* is not close to be established suggesting that its size can still growth as much genomes from independent strains are added Indicating an open pangenome.



Stenotrophomonas maltophilia

Strains	D457	K279a	JV3	R551
Genome size (bp)	4.769.156	4.851.126	4.544.477	4.573.969
Core genome size (bp)	2.764.114	2.764.471	2.761.327	2.764.543
% Genome	57,96	56,99	60,76	60,44
Total genes	4254	4354	4040	4023
Core genes	2742	2742	2742	2742
% gene	64,46	62,98	67,87	68,16
Pangenome contribution	4254	896	385	456

Figure 21: A) Accumulation curves of genes that contribute to the pangenome (A1) and the core genome (A2) of the four complete genomes of *S. maltophilia;* B) (B1) Circular representation of the pangenome from the for complete genomes of *S. maltophilia* indicating the amount of contribution of each added genome mapping the four genomes over the total pangenome. From the inner to the most extern blue ring: 1- S. maltophilia D457; 2- *S. maltophilia* K279a; 3- *S. maltophilia* JV3; 4 - *S. maltophilia* R551-3; (B2) Summary of the data obtained after calculate the core genome and pangenome of the four complete genomes of S. maltophilia. For each strain there was calculated the proportions of the genome that harbor the core genome and the pangenome.

Genome sequencing of 20 new *S. maltophilia* strains and general features

In this study were performed a detailed and comprehensive comparative analyses of 24 sequenced genomes of *S. maltophilia* (one complete genome sequenced along this study, three strains publicly available and 20 new strains, ten clinical isolates and ten environmental isolates sequenced exclusively for this study) collected between years 1993 and 2000. In total we analyzed the genomes of 12 clinical and 12 environmental isolates collected from different habitats.

The genomes of the 20 isolates of *S. maltophilia* sequenced using Illumina technology (California, USA) generated a total of 86.971.864 paired-end reads ranging from 35 to 336 bp in length. The average of G+C% for all strains after assembling was of 66.1%, in line with the described G+C% of the previously sequenced *S. maltophilia* strains. After sequencing, the sequences generated for each strain ranged between 3.353.214 to 6.492.918. All reads were submitted to filtering and trimming of contaminant sequences and further assembled *de novo* generating a total of 3.405 large contigs (>500 bp). The minimal number of contigs was obtained for the clinical strain FL2 with 108 contigs, which also harbors the largest one with 290.826 bp.

The largest number of bases assembled was generated for the clinical strain FL10, with a total consensus of 5.207.561 bp distributed in 359 contigs, harboring the maximum values for the number of genes and coding sequences (CDS), 4789 and 4728 respectively. The lowest number of genes (4066) and CDS (4007) were assigned to the clinical strain FL3. Detailed information about the general features of sequencing and assembling is presented at Appendix 1. For all new sequenced strains the CDS obtained were classified in subsystems using the RAST server (http://rast.nmpdr.org/) (Aziz et al., 2008). This classification did not show relevant differences in the composition of the new genomes of *S. maltophilia* at least in respect with the proportion of the distribution of genes assigned to subsystems in the genome (Figure 22).

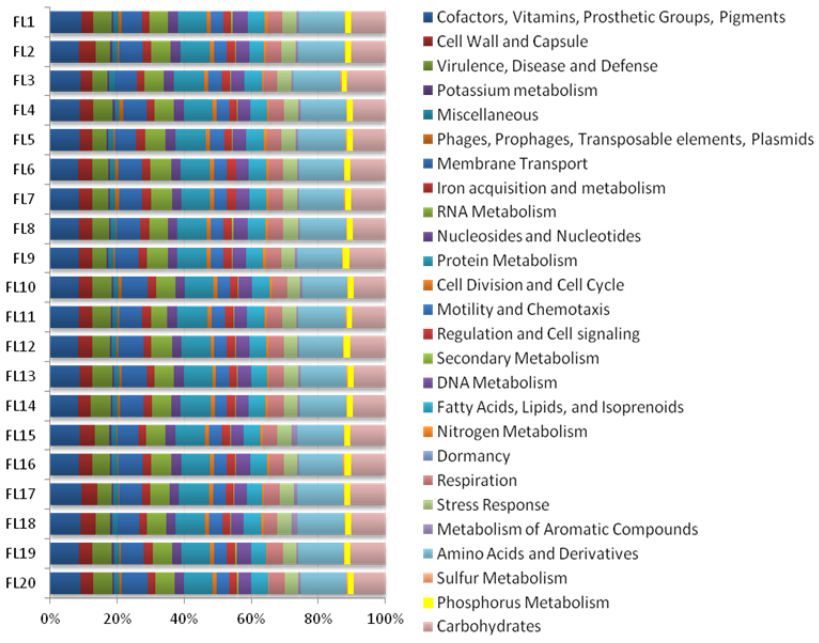


Figure 22: Distribution of subsystems identified from the functional annotation performed using the RAST Server. A homogeneous distribution of the subsystems along the clinical and environmental groups can be found.

Determination of the pangenome and core genome of 24 strains of *S. maltophilia*

At the beginning of this study, the parameters to perform the analysis of the core genome for the 24 strains were established by an array of settings using different combinations of coverage and identity (coverage 85, 90, 95 and 98%; identity 70, 75, 80, 85, 90, 95 and 98%) to generate the clusters of homologue genes (Figure 23). All clusters were generated using the algorithm COG triangles (Kristensen et al., 2010) embedded into the GET_HOMOLOGUES package (Contreras-Moreira and Vinuesa, 2013). Once the parameters have been established, all genes were clustered with at least 90% of coverage and 95% of identity, which correspond to 1875 clusters of genes shared by all 24 isolates. The genes with <90% in coverage and <95% identity were assigned as accessory genome corresponding, in average, to 2.350 genes for all strains indicating that more than a half of the detected genes for *S. maltophilia* are strain-specific (Table 5). The number of genes grouped in the core genome decreased markedly with the increasing of coverage settings (between 85-98%). However, no changes were presented at the core genome size when the coverage was maintained and the identity changed. This indicates that the different alleles of the different gene are correctly identified in our approach.

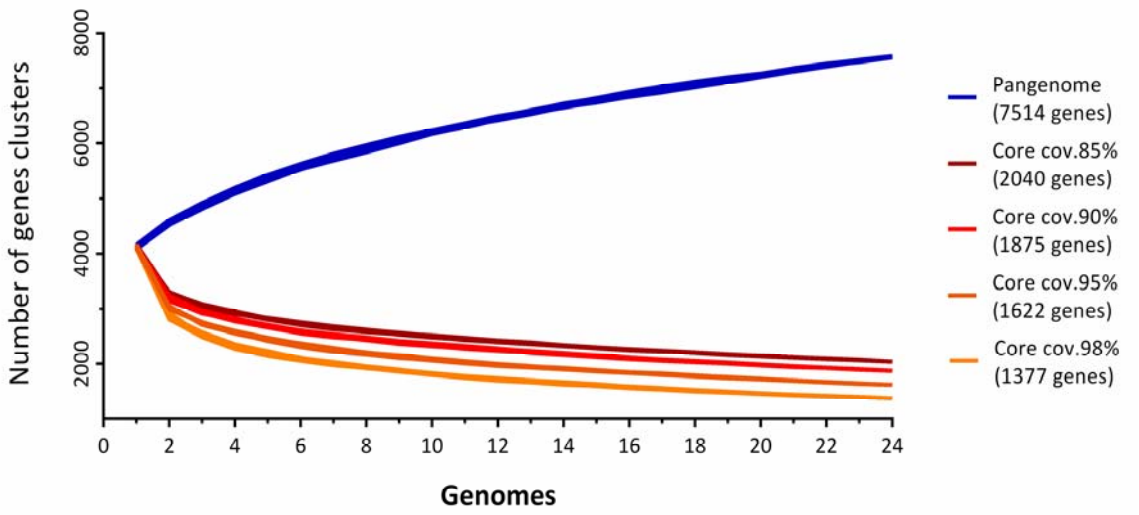


Figure 23: Array with different sets of coverage (85-98%), maintaining the value for the identity at 95%, previously performed to determine the core and pangenome for the 24 strains of *S. maltophilia*. The pangenome was constant with all settings while the core genome varied with the changes at the coverage.

A total of 83.793 genes were obtained from the 20 newly sequenced genomes plus 16.671 genes from the four reference strains D457, K279a, R551-3 and JV3, totaling 100464 coding sequences. All genes were grouped generating a total of 12214 genes clusters, and 1875 orthologous genes, which represented 15.35% of the total number of genes found and identified in all strains corresponding to the *S. maltophilia* core genome. For the 24 isolates this core genome represents, in average, to 44,86% (range: 41,27-48,32%) of the total genome for each strain. An overview at the proportion of each component of the pangenome for all strains demonstrated a homogeneous distribution of the core genome, the soft, the shell and the cloud genomes (Figure 24).

However, considering that not all genomes analyzed were closed (draft genomes) and some genes that certainly would be present in all genomes may be fragmented, the 'softcore' genome concept was considered at the time to determine the genes present in all strains extending the number of genes that form part of the core genome (Kaas et al., 2012). For this, genes shared by *n*-1 and *n*-2 strains were included in the core genome passing them from 1875 to 2585 genes (Figure 25). The remaining genes, present in several genomes and shared by 3 to 21 strains were named as the 'shell genome' and represented 3012 gene clusters indicating the vast number strain specific genes of *S. maltophilia*.

Table 5: Percentage, by strain, for the core and accessory genome of *S. maltophilia* using a minimum coverage ≥ 90% and identity ≥ 95% to generate the gene clusters resulting in a core genome with 1861 genes. *Complete available genomes of *S. maltophilia*.

Isolate	Genes	Core genes (%)	Accessory genes	Accessory genes (%)
D457*	4254	44,08	2379	55,92
K279a*	4354	43,06	2479	56,94
JV3*	4040	46,41	2165	53,59
R551-3*	4023	46,61	2148	53,39
FL1	4525	41,44	2650	58,56
FL2	4031	46,51	2156	53,49
FL3	3880	48,32	2005	51,68
FL4	4333	43,27	2458	56,73
FL5	3946	47,52	2071	52,48
FL6	4169	44,97	2294	55,03
FL7	4167	45,00	2292	55,00
FL8	4287	43,74	2412	56,26
FL9	4079	45,97	2204	54,03
FL10	4543	41,27	2668	58,73
FL11	4126	45,44	2251	54,56
FL12	4278	43,83	2403	56,17
FL13	4222	44,41	2347	55,59
FL14	4183	44,82	2308	55,18
FL15	4064	46,14	2189	53,86
FL16	4284	43,77	2409	56,23
FL17	4078	45,98	2203	54,02
FL18	4058	46,21	2183	53,79
FL19	4221	44,42	2346	55,58
FL20	4319	43,41	2444	56,59

With the intention to analyze whether or not clinical and environmental strains form different lineages in the evolution of *S. maltophilia*, we investigated the core and pangenome of these two groups separately. The comparison of clinical and environmental strains, with respect to their number of clusters, demonstrated that the pangenome of clinical strains was represented by 9542 genes in comparison with the 8747 genes present in the pangenome of the environmental strains (Figure 25). Despite their 'pangenomic' differences, both groups presented a similar core genome size, corresponding to 2205 genes for the environmental strains and 2214 genes to the clinical strains. The clinical strains FL1 and FL10 presented the higher values for the accessory genome presenting 2709 and 2727 genes, respectively. It is worth mentioning that the clinical strain FL3 was the isolate with the smaller accessory genome presenting 2064 genes.

The softcore and shell did not presented relevant differences at least corresponding to the number of genes presenting 2897 and 3005 genes for the environmental and clinical isolates softcore, respectively. The shell genes of the environmental and clinical strains grouped 1544 and 1340 genes, respectively. Analyzing the cloud genome for each group of isolates there was found a remarkably difference at the number of genes shared by only two strains and genes exclusively present in one isolate. Those genes not shared by all strains but only with a few isolates represented the cloud genome of both groups. The cloud genome of the environmental strains was composed by 4306 genes meanwhile the clinical strains presented a total of 5197 genes.

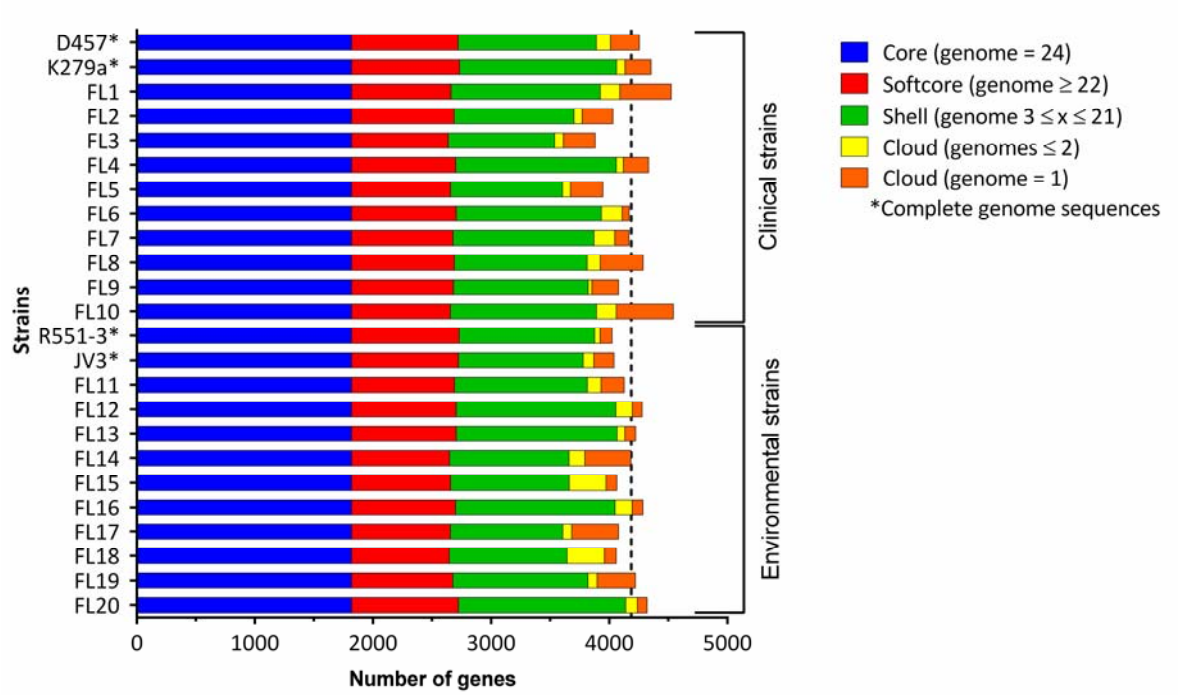


Figure 24: Representation of the core and accessory genome distribution of clinical an environmental strains of *S. maltophilia*. The soft core was defined as the *n-2* of the total genome collection. Dot line represents the meaning value of the genome length of the strains.

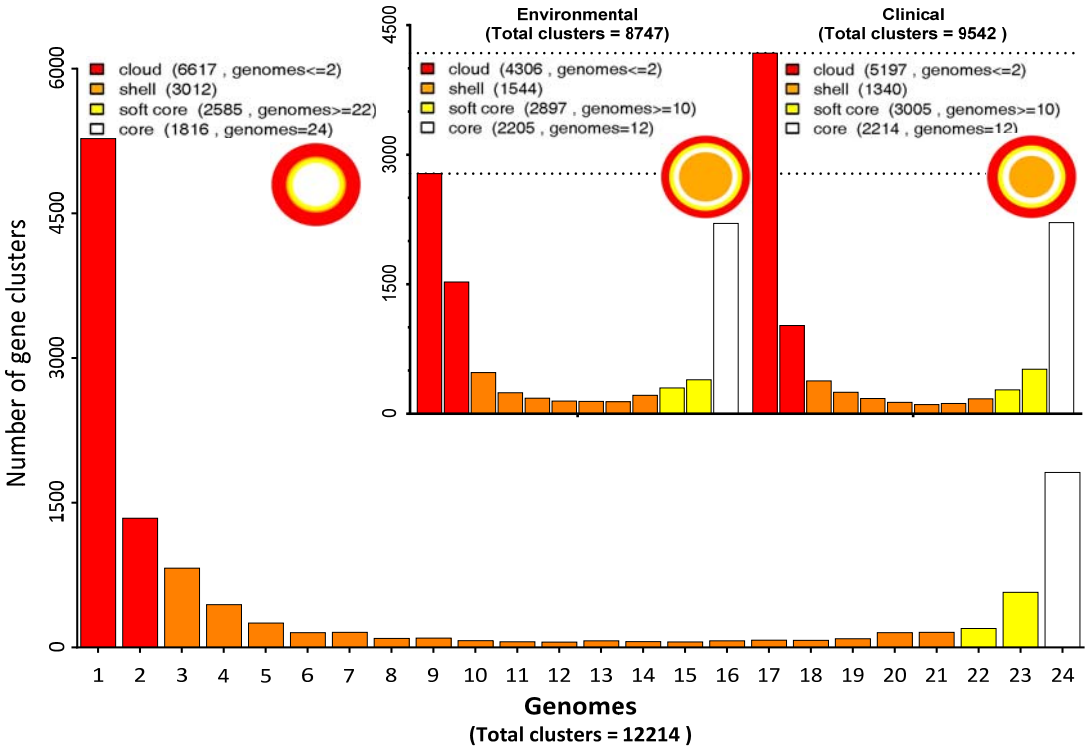


Figure 25: Distribution of genes of the 24 genomes used to calculate the core genome, soft core, shell and cloud genome. The major graphic shows all genomes used in this study. Minor charts show the distribution of the clusters of environmental and clinical strains. For all charts: White bars – core genomes; Yellow bars – soft core genes; Orange bars – shell genes; Red bars – cloud genes.

Calculating the pangenome size, the results suggest that the number of gene families will asymptotically increase with respect to the number of analyzed strains, which may indicate that *S. maltophilia* has an "open pangenome" based on the 24 genomes examined (Tettelin et al., 2008) (Figure 26).

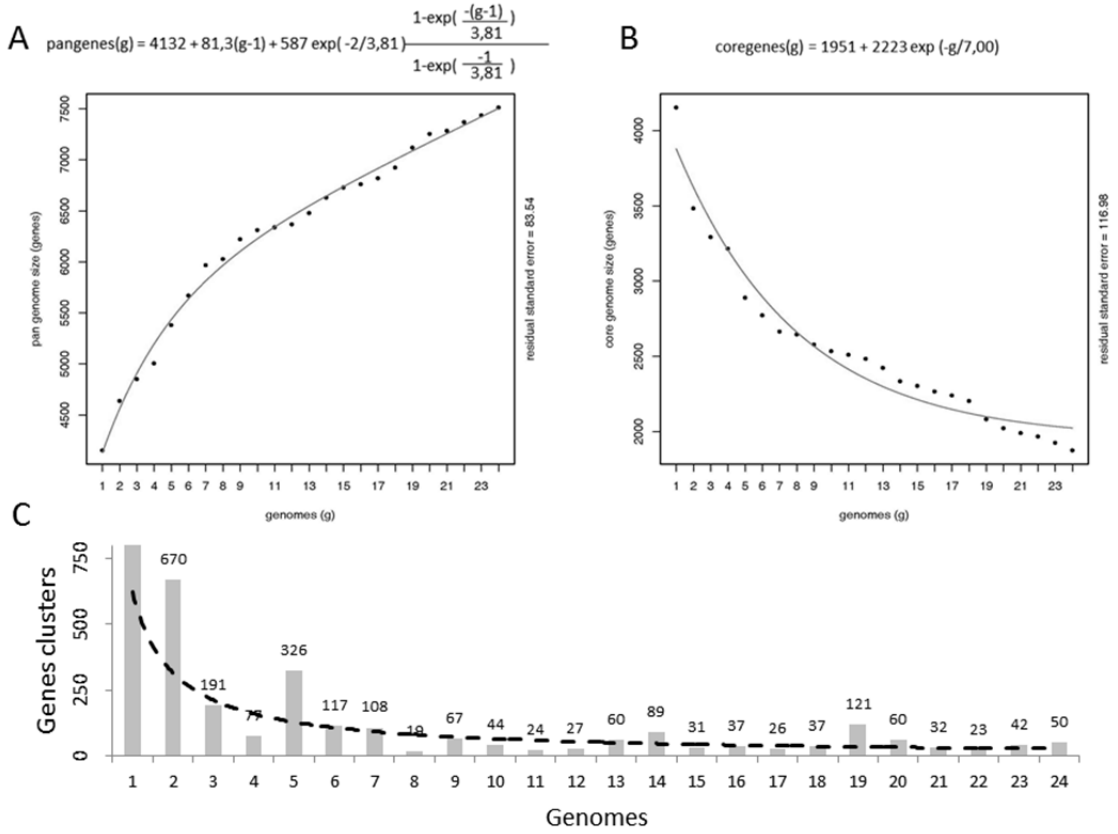


Figure 26: Pangenome and core genome curves representing the contributions of 24 strains analyzed in this study. The vertical bars represent the number of genes added to the pangenome after the inclusion of one strain. A - black points represent the amount of genes included at the pangenome as a function of the number of genomes added in the set of strains. B - black points represent the amount of shared orthologous plotted as a function of the number of genomes added. Solid curves (red) represent the power of tendency of the points for the pan and core genome. C – Representation of the contribution of genes before adding a new genome to the analysis to calculate the pangenome.

The analysis of more strains is needed to fully ascertain how open is the pangenome of *S. maltophilia*. Analyzing the trends, new gene families will continue to be found as many genomes are included in the analysis. Consistent with this situation, the size of the core genome is inversely dependent of the number of strains included in the analysis as well of the identity and coverage settings used to determine the cluster families. Nevertheless, if the soft-core in included as forming part of the core-genome, the size of the core genome would be in the range of 2000 genes. The sequenced of more closed genomes are required to get a more accurate estimation of the sizes of the core and the pangenome of *S. maltophilia*

Clinical and environmental strains are not two different evolutionary lineages

The phylogenetic relationships among the 24 *S. maltophilia* genomes were examined using the gene composition to distinguish whether or not clinical and

environmental strains constitute two independent evolutionary lineages. Five major clusters where obtained, none of them justifying a clear separation of the clinical and environmental strains in base of their genes composition, with at least one distinct strain present in each of the groups independently of their origin (Figure 27).

Six strains were clustered at the major branch "A", which is composed by four clinical strains (FL1, FL5, FL8 and FL10), isolated from urine and blood, and two environmental strains (FL11 and FL14). The strain FL11 has no information concerning the habitat from where it was collected but FL14 also came from an aquatic niche. In common, all of these strains had their origin from aquatic and saline habitats. The cluster "B", with four isolates, is composed by the strains which have their genome closed with D457 and K279a as clinical strains and R551-3 and JV3, the environmental isolates. The branch "C", with five isolates we found the most diverse group with respect to their origin, containing clinical and environmental isolates obtained from urine (FL6), sputum (FL7), pus (FL9), rhizosphere (FL12) and eye-care solution (FL16). Cluster "D" with four strains presented three environmental isolates and two clinical, the same occurring with cluster "E" which was formed by only one clinical strain originated from a blood sample (FL4) and three environmental strains isolated from sewage. The cluster "E" was formed by three environmental strains sampled from the rhizosphere, FL15, FL17 and FL18, and two clinical isolates, FL2 and FL3. Both clinical strains were obtained from the sputum and respiratory secretion, respectively.

Despite the formation of five related clusters, these results indicate that independently on whether they have a clinical or an environmental origin, these strains present a similar composition in genes. The results showed in this Thesis support the idea that both clinical and environmental isolates are closely related and all of them likely may or may not present a pathogenic behavior depending on the circumstances in which they are presented.

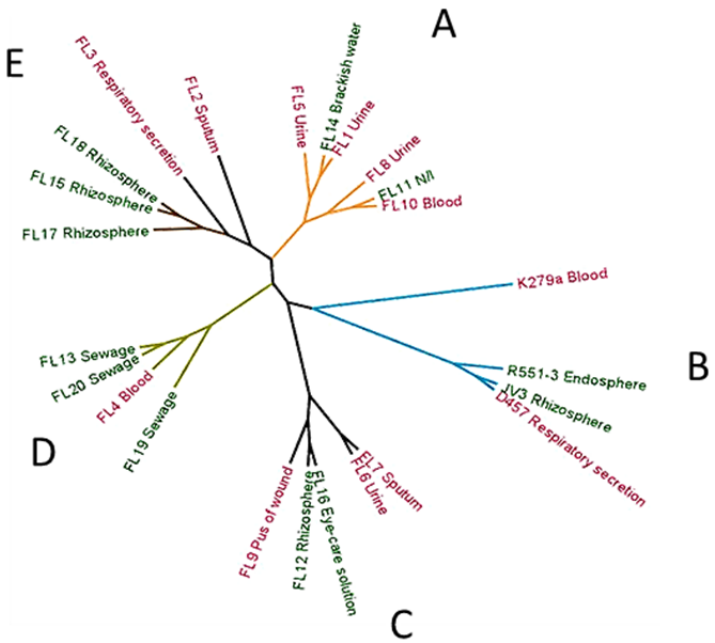


Figure 27: Cladogram calculated by the CDS composition of 12 clinical (red) and 12 environmental (green) strains of *S. maltophilia*. The cladogram was constructed using the software FigTree v.1.4.2 with midpoint root. The presence of both groups in all branches represent that, independently of their origin, S. maltophilia strains are composed by similar set of genes suggesting the same pathogenic potential.

Functional-based genomics comparison between clinical and environmental strains of *S. maltophilia*

Different origins may determine distinct behaviors, and even closely related strains may present specific sets of genes, including resistance or virulence determinants that have been acquired from other bacteria. The major problem to use *S. maltophilia* strains as a biotechnological resource is the fact that this opportunistic pathogen presents multiple mechanisms of resistance to many antimicrobial compounds as well as virulence factors responsible for host infections and niche colonization. To have further insight on those aspects that are fundamental for *S. maltophilia* infection, the 12 clinical and the 12 environmental strains of *S. maltophilia* were compared in respect to their resistance genes, virulence factors, secretion systems, mechanisms of cell-to-cell signaling communication and antibiotic susceptibility.

In the scope of this work, thirty-four resistance genes and putative resistance genes present in D457 strain and at the publically available genomes of *S. maltophilia* were used as a platform to detect the presence of these genes into the other 23 genomes studied. Genes grouped at the same cluster as that of the subject genes were considered as homologous and in cases of doubt they were checked manually making a BLAST search to confirm their presence or absence. The strain that presented the highest number of resistance genes was the clinical strain K279a, followed by the environmental strains FL12 and FL15 (30 genes).

The resistance genes *bla1* and *bla2* that confer resistance to beta lactam antibiotics were detected in all strains of both groups. Meanwhile, bla_{AmpC} beta lactamase, which mediate resistance to cephalothin, cefazolin, cefoxitin and most penicillins, were detected in 21 out of 24 isolates, and was not detected neither the clinical strains FL2 and FL5 nor in the environmental isolate FL14. In addition, a putative beta-lactamase bla_{AmpC} was found in 13 isolates including the strains FL5 and FL14. Six putative beta-lactam resistance genes of *S. maltophilia* D457 (GI:504460542, GI:504460526, GI:754362052, GI:504458662, GI:504460856, GI:504460542) and two from *S. maltophilia* K279a (GI:501457908; GI:501456828) were recognized in most of the clinical and environmental strains. The fluoroquinolone resistance gene *Smqnr,* which confers low-level resistance to quinolones and fluoroquinolones (Zhang et al., 2012) was detected in all the isolates. Five putative metallo-beta-lactamases were also found in both groups of strains (Appendix 2).

The mechanisms used by all Gram-negative bacteria to secrete and transport effectors proteins through the membrane without compromising the periplasm were analyzed. On the basis of bioinformatics approaches, two signal-dependent secretion proteins, type2-secretion (T2SS) and type5-secretion (T5SS) and the non-signal dependent putative secretion protein system, type4-secretion (T4SS) were found in all genomes of *S. maltophilia* indicating to be constitutive genes present in the core genome of this species. Otherwise, two putative protein type1-secretion (T1SS) recently described at *S. maltophilia*

RA8, an environmental strain isolated from the effluent of the sewage treatment plant, and the clinical strain SKK35, collected from an ulcer swab (Adamek et al., 2014, 2011) were detect in some of the genomes studied in this work. The RA8 T1SS encoding genes (4243-44) were found sharing a gene cluster with eight other *S. maltophilia* strains - five clinical (D457, FL1, FL2, FL3, FL8) and three environmental (R551-3, FL15, FL19) - demonstrating to have a restricted range of distribution but also that this restricted range is not linked to the ecosystem where they were isolated (infected patient of natural environment). The presence of the putative T1SS proteins firstly described at the genome of the isolate SKK35 (1685-86) was analyzed using these sequences as query to search in our database created with all CDS observed in the 24 genomes of *S. maltophilia* studied, and found to be present just in the clinical isolate FL9 (stml-FL9_03981-82) with no homologous regions shared with our set of genomes. With 95,56% of identity for the CDS stml-FL9_03981 and 97,08% for stml-FL9_03982 respectively, the SKK35 (CDS 1685-86) putative leukotoxin secretion complex was recognize at the isolate FL9 (Camacho et al., 2009). This finding indicates a restricted distribution of this complex, occurring only at one strain of 24.

The recently discovered type6-secretion (T6SS) is generally encoded by a group of 13 conserved genes required for function with an average size of 20Kb (Cascales and Cambillau, 2012), and was previously found composing the genome of the clinical isolate *S. maltophilia* SKK35 (Adamek et al., 2014). A direct search of the coding sequences for those proteins that compound the complex permitted to discharge their presence in three environmental strains isolated from rhizosphere of plants, FL15, FL17 and FL18, while all the other 21 isolates presented the genes encoding T6SS.

Regarding the elements involved in movement and colonization of different niches, hosts or surfaces, adhesion and motility factors, such as pili and flagella, were found in all sequenced genomes. The chemotaxis genes, *cheRB* and *cheBDR*, the flagella motor proteins, as well the assembly regulator *fliSD* and the flagella components *flqLJIHGFEDC* were observed in all studied genomes.

The gene *SmpA*, an outer membrane protein which in *Escherichia coli* and its homologous *OmpA* in *Pseudomonas fluorescens* is responsible for the assembly of outer membrane β-barrel protein at the multicomponent *YaeT* complex (OMPs), was found in nine organisms (K279a, FL3, FL6, FL7, FL9, FL12, FL15, FL16 and FL18).

The detection of genetic differences which possibly contribute to virulence of the clinical and environmental isolates of *S. maltophilia* widely distributed among all genomes were performed with the newly sequenced strains comparing with the complete genomes of *S. maltophilia* D457, K279a and R551-3. The strain JV3 did not present any specific gene which was not present in the other three strains and for this reason it was not considered when the list of genes were compiled, using available information on genes encoding virulence determinants in this species. Forty-five genes were searched into the 24 genomes and all isolates harbored genes encoding for hemolysin III, T5SS hemaglutinin, metalloprotease, phospholipase and the siderophore enterobactin synthetase.

Investigating 32 virulence factors of all 24 strains, 21 were found in all genomes. The other virulence factors analyzed presented a distribution from 1 to 23 genomes highlighting the genes *lkt*B and *hly*B present only at the genome of the clinical isolate FL9 obtained from one wound. Since the presence of specific genes may differentiate one organism from another, the average similarity distance of each isolate was calculated based on a presence/absence matrix for all strategic genes (Appendix 2) used as features to check the correlation of the isolates (Figure 28A-B). As shown, the environmental and clinical isolates did not cluster independently.

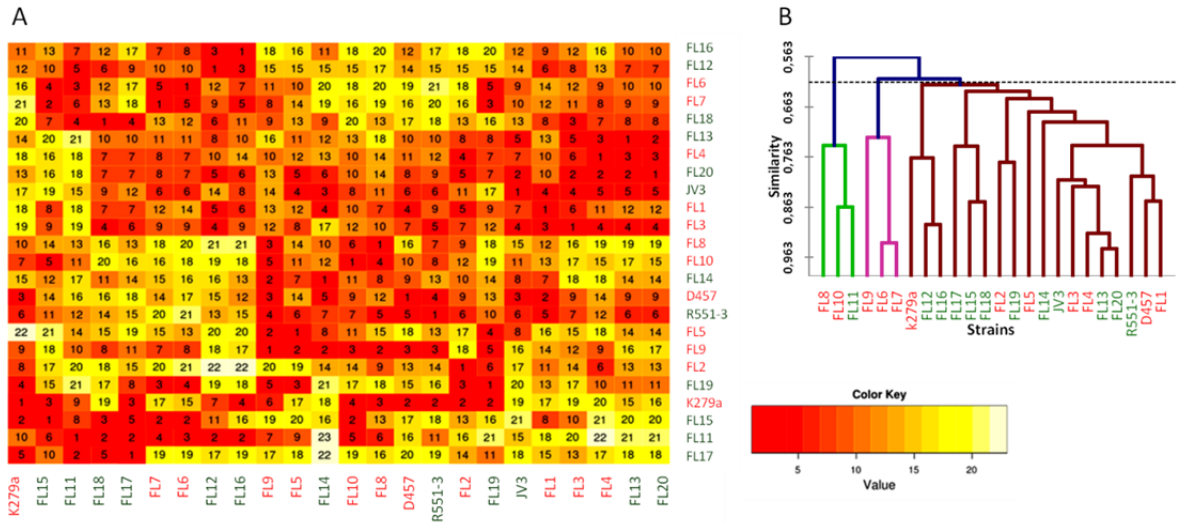


Figure 28: A) Heatmap created from a presence/absence matrix for all strategic genes representing the proximity of each isolate independent of their origins. The matrix served to calculate the average similarity distance based on the genes present; B) Dendogram representing the distribution of the clusters formed from the strategic genes analyzed in all genomes.

Antibiotic susceptibility

The Minimal Inhibitory Concentrations (MIC) of 12 different antibiotics from different families was determined for all the isolates using antibiogram strip-tests to check their susceptibility to antimicrobials frequently used in clinical treatments (Table 6). The quartile for the MIC values in each group was determined to plot the distribution of isolates with respect to the concentrations of antibiotics. The antibiogram strip-tests for each separated group demonstrated that the clinical isolates, as a group have a trend towards higher levels of resistance (Figure 29).

The environmental strain FL15 demonstrated to be the unique isolate susceptible to IMI and ETP, while the others grew over the maximum value of this strip-test (32 µg/ml) which fits with the available information that describes *S. maltophilia* as resistant to these antibiotics (Howe et al., 1997). On the other hand, the same isolate FL15 presented the highest level of resistance to CL (>256 µg/ml), with the second more resistant isolate to this antibiotic (FL8) presenting a MIC of 128 µg/ml and the others ranging between 3-32 µg/ml. This may mean that despite the overall higher resistance of clinical isolates as compared with environmental ones, all *S. maltophilia* isolates present similar chances to acquire resistance independently of their origin.

Considering only the antibiotic concentration ranges where the values did not exceeded the maximum concentration of the strip tests, the clinical strains demonstrated to group those strains presenting higher MIC values for the antibiotics SXT, TGC, GAT, MXF and NA in comparison with the isolates with environmental origin. Distinct of the clinical strains, the environmental group demonstrated a restricted range of resistance to the antibiotics used in this study occasionally exceeding the range presented by the clinical strains (Figure 29 B, K). The multiple-antibiotic-resistance pattern of the environmental strains, showing in several cases closely MIC values to those shown by the clinical strains, could be explained by the presence of an intrinsic resistome linked to the core genome of these strains.

Table 6: Minimal Inhibitory Concentrations (MICs) of 20 new sequenced strains and S. maltophilia D457.

Strains	SXT	TGC	CAZ*	PM*	CN	GAT	CS	CL*	IMI*	ETP*	MXF	NA
FL1	0,75	3	1	4	2	1	4	24	>32	>32	0,75	8
FL2	1	2	>256	>256	1	128	6	32	>32	>32	0,25	4
FL3	0,38	2	4	12	24	0,25	4	16	>32	>32	0,13	8
FL4	0,38	0,75	>256	48	2	0,06	24	3	>32	>32	0,09	6
FL5	0,19	0,75	1,5	6	0,38	0,13	24	4	>32	>32	0,19	6
FL6	0,25	0,75	96	64	24	0,13	24	6	>32	>32	0,06	3
FL7	0,38	0,5	128	64	16	0,13	24	8	>32	>32	0,13	4
FL8	0,75	6	192	96	32	3	48	128	>32	>32	3	48
FL9	0,5	0,75	256	64	1	0,13	256	8	>32	>32	0,13	6
FL10	0,64	0,09	12	16	0,5	0,05	3	8	>32	>32	0,06	4
FL11	0,19	0,75	16	64	2	0,13	12	6	>32	>32	0,13	3
FL12	0,09	0,19	16	32	4	96	12	16	>32	>32	0,13	0,8
FL13	0,5	0,75	32	48	2	0,13	32	6	>32	>32	0,03	3
FL14	0,5	0,19	32	48	4	0,02	8	16	>32	>32	0,03	2
FL15	0,02	0,05	64	>256	0,09	0,02	16	>256	0,06	0,2	0,03	8
FL16	0,19	0,25	24	64	4	0,19	8	6	>32	>32	0,09	2
FL17	0,09	0,19	16	32	6	0,09	48	8	>32	>32	0,03	2
FL18	0,02	0,75	24	32	4	96	8	6	>32	>32	0,03	2
FL19	0,19	0,5	>256	64	1	0,09	16	16	>32	>32	0,06	3
FL20	0,13	0,38	32	48	6	0,06	48	8	>32	>32	0,06	6
D457	0,13	1,5	1,5	16	6	0,5	32	12	32	32	0,25	8
MIC50	0,25	0,75	32	48	2	0,13	16	8	32	32	0,09	4
MIC90	0,75	2	256	96	24	96	48	32	32	32	0,25	8

The values for the MIC50 and MIC90 are also included. (SXT- Trimethoprim/Sulfamethoxazole; TGC- Tigecyclin; CAZ- Ceftazidime; PM- Cefepime; CN- Gentamicin; GAT- Gatifloxacin; CS- Colistin; CL- Chloramphenicol; IMI- Imipenem; ETP- Ertapenem; MXF- Moxifloxacin; NA- Nalidixic Acid). *At least one strain grew over the maximum value of the strip-test indicated by ">" when this value exceed the limit.

To further analyze whether clinical and environmental isolates behave differently in terms of antibiotic resistance, the values of the strip tests were normalized to either the MIC50 or the MIC values for *S. maltophilia* D457, and the isolates were clustered (Figure 30-A, -B). The normalization of the MICs with the MIC50 of all 24 isolates highlighted seven environmental strains (FL11, FL13, FL14, FL16, FL17 and FL20) grouped in one branch where they indicated to be less susceptible than the other strains to the carbapenems imipenem and ertapenem, and the cefalosporins ceftazidime and cefepime (Figure 30-A). In other small group, the strains FL12 and FL18 demonstrated to be the most resistant strains to gatifloxacin in comparison with the other isolates. On the other hand, these strains were the most susceptible to the remaining antibiotics.

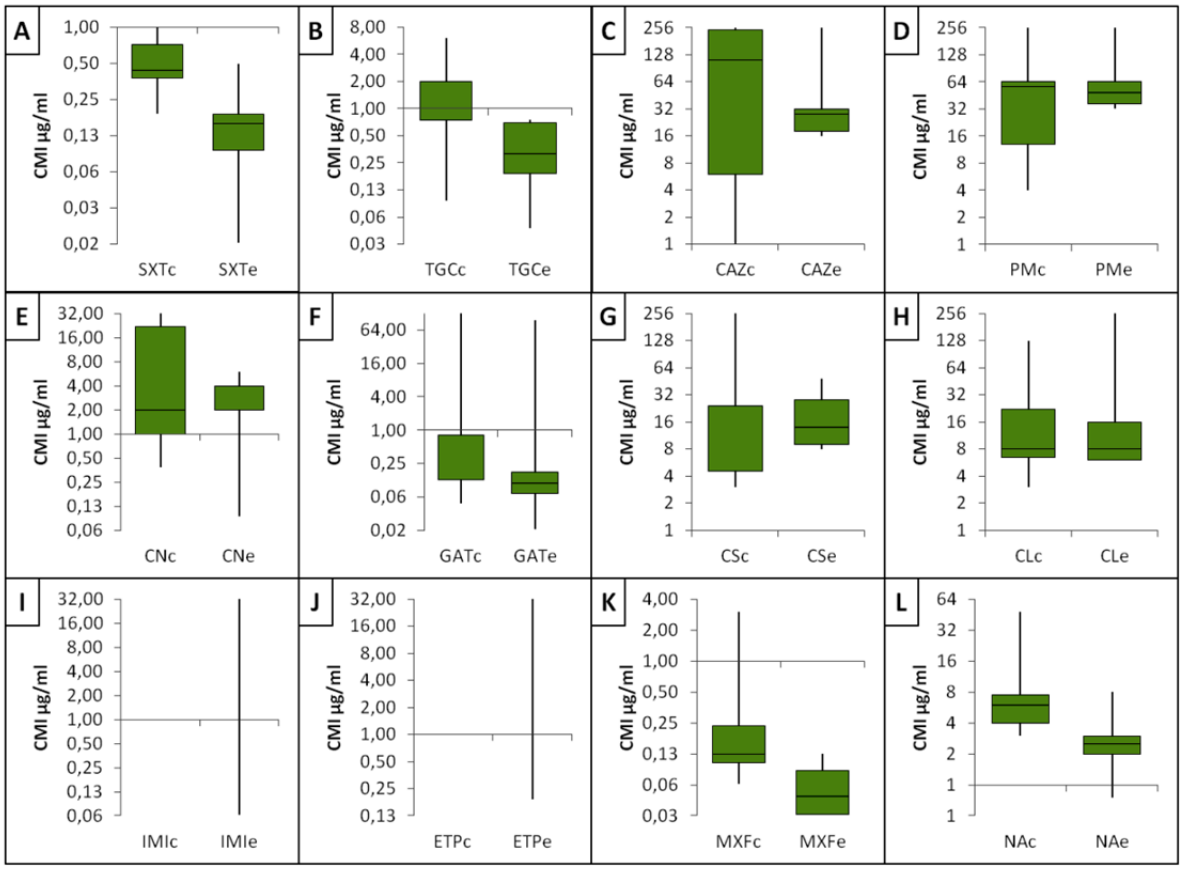


Figure 29: Boxplot charts representing the Minimal Inhibitory Concentrations (MICs) for all clinical (c) and environmental (e) isolates obtained using antibiogram strip-test of 12 antibiotics from different families (SXT- Trimethoprim/Sulfamethoxazole; TGC- Tigecyclin; CAZ- Ceftazidime; PM- Cefepime; CN- Gentamicin; GAT- Gatifloxacin; CS- Colistin; CL- Chloramphenicol; IMI- Imipenem; ETP- Ertapenem; MXF- Moxifloxacin; NA- Nalidixic Acid). There was calculated the quartile for the MIC values in each group to plot the distribution of isolates with respect to the concentrations of antibiotics.

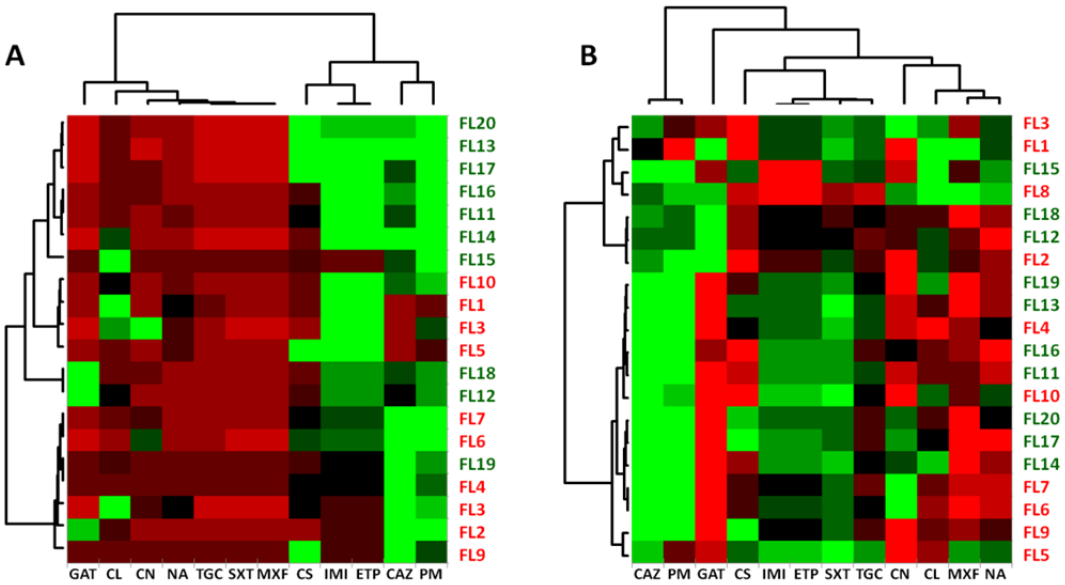


Figure 30: Heatmap plotting the relative susceptibility of 20 *S. maltophilia* strains **A**) obtained from the normalization of the MICs by the MIC50 and **B**) the values normalized with the MICs of *S.maltophilia* D457 for the same antibiotics. To all strains, the relative susceptibility was expressed in logarithmic scale (log2). (SXT- Trimethoprim/Sulfamethoxazole; TGC- Tigecyclin; CAZ- Ceftazidime; PM- Cefepime; CN- Gentamicin; GAT- Gatifloxacin; CS- Colistin; CL- Chloramphenicol; IMI- Imipenem; ETP- Ertapenem; MXF- Moxifloxacin; NA- Nalidixic Acid). Green plots represent the increasing value (more resistant) while red plots represent the decreasing MICs (less resistant).

Quorum sensing regulons

The quorum-sensing (QS) system is responsible for the synchronization of particular bacterial behaviors on a population scale. In the case of *S. maltophilia* this process depends on the Diffusible Signal Factor QS (DSF-QS) system, which is based mainly on the fatty acid DSF (cis-11-methyl-2-dodecenoic acid) (Fouhy et al., 2007).

In a previous study there were found two variants for the *rpfF* gene, which is responsible for the regulation of DSDF-AQ (Fouhy et al., 2007; Pol Huedo et al., 2014). Since these variants are markers of two different phylogenetic branches, each one presenting differences in terms of virulence, we analyzed their presence in the 24 studied genomes. Using the available sequences (GenBank accession numbers KJ149475 to KJ149552) a direct search were performed for the corresponding DNA region of *rpfF* gene. It demonstrated that the 24 strains of *S. maltophilia* harbor this gene with different length and variable residues along the sequence. This variability revealed that *rpfF* is distributed into two distinct groups separating the strains with respect to the variants. Each group was both composed by 12 strains (Figure 31). Analyzing the origin of the strains into each group formed, there was no preferential distribution of the isolates, with clinical and environmental strains presenting an equitable distribution. The cluster with the RpfF-v1 variant presented five environmental strains (R551-3, FL11, FL12, FL16 and FL19) and seven clinical isolates (K279a, FL6, FL7, FL8, FL9 and FL10); on the other hand the cluster containing the RpfF-v2 variant presented seven environmental strains (JV3, FL13, FL14, FL15, FL17, FL18 and FL20) and five isolates representing the clinical (D457, FL1, Fl3, FL4 and FL5).

Figure 31: Comparative analysis of 24 *S. maltophilia* strains based on the first 108 amino acids of RpfF. The colored bars represent the amino acids residues indicating a clear separation between the two variants of RpfF.

Metagenomic analysis of the effect of triclosan on antibiotic resistance

It has been proposed that WWTP can be hotspots for the acquisition and evolution of antibiotic resistance, because they can contain bacterial pathogens released from human samples, together with antibiotics and other selective compounds as biocides. To have a deeper insight on the role that biocides may have in the selection of antibiotic resistance in WWTP a study has been performed in collaboration with Dr. Ülkü Yetiş from Middle East Technical University (Turkey). Activated sludge from a WWTP was incubated in the presence of triclosan following two different regimens as described in Methods..

Metagenomic DNA was obtained and treated as described in Materials and Methods and sequenced using Illumina technology. A total of 147.863.659 reads were generate, paired-end reads passing quality filtering of sediment samples. The reads were checked separately using FastQC software and filtered using the NGS QC Toolkit to clean contaminant sequences and low quality reads, producing roughly 129.747.442 reads distributed with all samples. For each sample the resulting reads after filtering were joined to overlap over each one and generate larger sequences (bireads). Each set of bireads were compared with the four customized databases: 1 – Antimicrobial resistance genes (1597 sequences); 2 – Antibiotic Resistance Cassettes (316 sequences); 3 – Integron DNA sequences (3286 sequences) and 4 – Plasmid genomes (839 sequences).

Antimicrobial resistance genes

In order to analyze the frequency and diversity of resistance genes, a BLASTp search with all overlapped reads was applied. Using this approach, we initially found a total of 340 genes with an expectation value (*e-value*) of 10^{-10} distributed at nine samples that were homologous to sequences found at the ARPCARD resistance genes database. Analyzing in deep all the putative resistance genes identified at both experiments, a manual curation of the results was performed to remove false-positive resistance genes and identify the putative location for each one; if it was present in plasmids or at the bacteria chromosome. After curation, only 268 genes remained for the further analysis with 90 present at the chromosome and 178 identified as putative plasmidic genes. Sequences which did not contribute directly to a resistance phenotype were discharged, but are only involved in resistance upon mutation.

The sample used as control for the first experiment (Exp1) presented a total of 108 curated resistance genes and the second sample, used as control for the second experiment (Exp2) revealed 126 resistance genes present in the metagenome. The diversity of resistance genes along both experiments suffered changes, ranging from 100 to 271 genes at the Exp1 and from 81 to a 100 in Exp2

The abundance of resistance genes was also analyzed. In Exp1, a peak of high abundance of resistance genes was observed at 100 ppb of triclosan with most reads corresponding to plasmid genes (Figure 32). This abundance does not correlate with diversity. Indeed, at Exp1 the sample of 1 ppm of triclosan presented the higher diversity of

resistance genes with 131 recognized as plasmid genes and 61 assigned as chromosomal genes. This increasing in diversity was not followed by the abundance of bireads, which reached the value of 2,70% of the total of sequences for this sample. At the highest triclosan concentration, a reduction at the abundance of resistance genes was observed; 0,06% of the overlapped reads were recognized as resistance genes at the level of identity established.

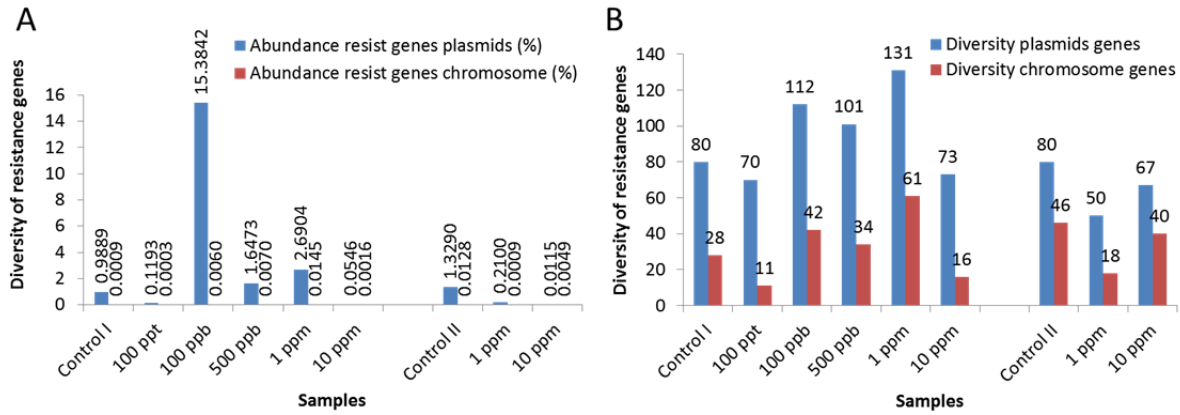


Figure 32: Distribution of diversity and abundance of resistance genes found at the nine microbiota samples sequenced to this study. The resistance genes were analyzed with an e-value of 10^{-10} and curate manually to eliminate false-positive genes at the samples, maintaining only real antimicrobial resistance genes, for which a plasmidic or chromosomal origin was inferred. In all samples the diversity of genes which proceed from plasmids was higher than the diversity of genes with chromosomal origin. The abundance of bireads followed the trend and the plasmid genes were higher than the chromosomal resistance genes.

In Exp2, where two different triclosan concentrations were applied without intermediate steps to reach the maximum concentration, the initial sample of this experimental regimen (Control II) presented 80 resistance genes belonging to plasmids and 46 to chromosomes. Differing to the situation observed when the activated sludge was submitted to sequential increased triclosan concentrations, the application of fixed amounts of the biocide reduced the abundance and diversity of resistance genes, both plasmid-encoded and chromosomal genes (Figure 32).

In general lines the plasmid genes were present in higher numbers when compared with the chromosomal genes. However this could be the consequence of a bias in the database used in this study. On the other hand abundance did not correlate with the diversity of genes identified using an e-value of 10^{-10}. The abundance of bireads found at 100 ppb suggests the presence of specifics taxa which grow better at this concentration of triclosan. In this sample the abundance was shifted by the beta-lactamase genes POX7-1_p1, TEM-1, TEM-63, TEM-67, AB57_0283 and TEM-168 with 398959, 198903, 133746, 129945, 88917 and 65850 bireads respectively.

A close view at the resistance genes maintained during all the experiment reveals that 35 plasmidic genes were shared by all six samples at Exp1 changing their abundance according to the concentration of triclosan applied (Appendix 6). At the second regimen, 23 genes were present at 1 pmm and 10 ppm of triclosan, from which 11 presented a chromosomal origin (Figure 33).

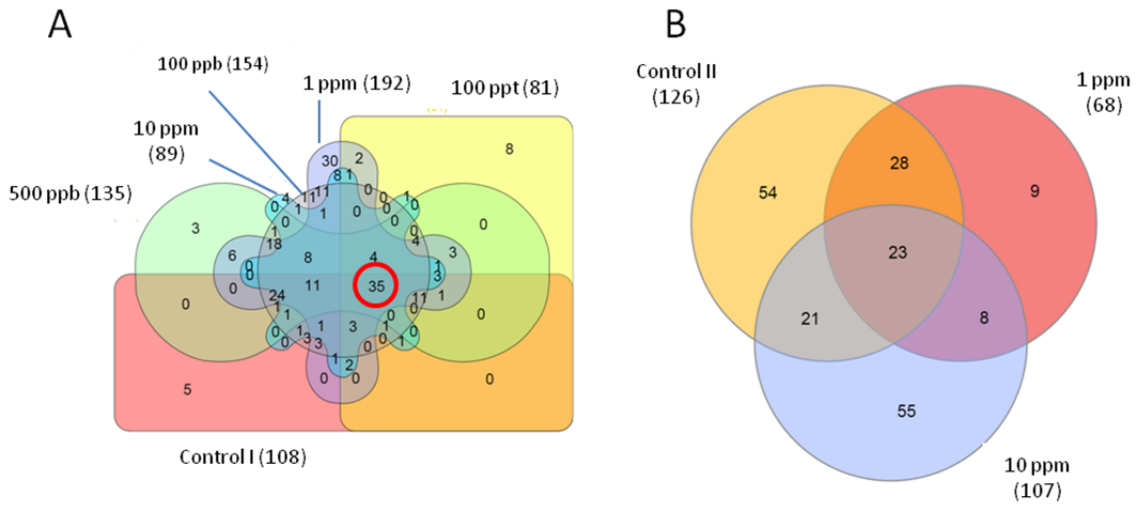


Figure 33: Venn diagram with the distribution of resistance genes along the two regimen of treatment with triclosan. A - Diagram showing the six samples and how the genes are distributed. The red circle highlights the number of genes present at all samples along the experiment; B – Venn diagram with the distribution of genes present at the three samples from the second regimen of biocide treatment. All samples shared 23 genes which were identified at low and high concentrations of triclosan.

The major part of the antibiotic resistance genes found in all samples was recognized as belonging to the most commonly type of beta-lactamase encountered in plasmids, the TEM type. At least 57 TEM beta-lactamase genes were initially recognized into the samples. To identify which TEM was actually present in the samples, an alignment approach was performed to map the filtered reads against all known TEM genes previously identified obtaining the entire reconstructed gene sequence. The sequences submitted to an alignment search against a type TEM beta-lactamase database whit 100% of identity and 100% of coverage with no gaps were assigned as the correspondent TEM. The beta-lactamases identified with the first approach that not reached the criteria of 100% of coverage and identity was assigned as TEM-like. Only 21 TEM were completely confirmed using this approach (Table 7).

Table 7: Summary of the identification of the beta-lactamase TEM genes present at the nine metagenomic samples.

Samples	Matches	Total reads	% reads	Reads coverage	Estimated	High ID
Control I	91.747	16.277.158	0,564%	175,88	37	12
100 ppt	9.913	11.118.858	0,089%	50,2	16	0
100 ppb	1.198.424	13.403.780	8,941%	2297,34	53	20
500 ppb	151.571	12.195.468	1,243%	737,76	42	4
1 ppm	200.672	11.538.794	1,739%	945,88	52	8
10 ppm	4.778	18.432.952	0,026%	9,16	11	0
Control II	131.497	16.743.438	0,785%	252,08	43	8
1 ppm	18.499	11.836.620	0,156%	95,68	25	2
10 ppm	820	18.200.374	0,005%	1,57	8	0

Antibiotic Resistance Cassettes and Integrons

Resistance genes may be fixed at the chromosome or inserted in mobile elements as plasmids ready to be transferred to another host. Gene cassettes are another type of mobile element but in this case the transfer of DNA molecules by recombination occurs in

the same cell. Those gene cassettes do not have their own machinery to move into the genome and for this are usually found integrated into elements called integrons.

To understand the influence of triclosan at the composition and abundance of resistance genes and integrons in a simulated environment, the bireads from each sample were used as query sequences against two customized databases containing cassettes genes and integrons sequences. The total of diversity of cassettes genes found in both groups summed 291 registers. Samples submitted to increasing concentrations of triclosan presented an initial diversity of 134 cassettes and a final consensus diversity of 217 cassettes genes. The diversity of cassette genes at Exp1 suffered a reduction to 69 (abundance of 0.02%) after one week since the first exposure of triclosan. At 100ppt of triclosan the diversity and abundance of theses mobile elements grew by the appearance cassettes not recognized at the control sample. To Exp2, the initial diversity was of 111 cassettes at the original sample decreasing in value at 1ppm with 99 cassettes and an abundance of 0,80% of the total bireads identified with an *e-value* of 10^{-10}. The subsample submitted to the concentration of 10ppm of triclosan presented the higher diversity and lower abundance, with 153 and 0,14% respectively.

Despite the variations at the concentration, diversity and abundance, 43 resistance genes cassettes were recognized in all samples of Exp1 and 69 at Exp2, changing their abundance with respect to each concentration of triclosan the samples were submitted (Figure 34). In all samples of both regimen the appearance of "new" individuals did not compromised the final abundance.

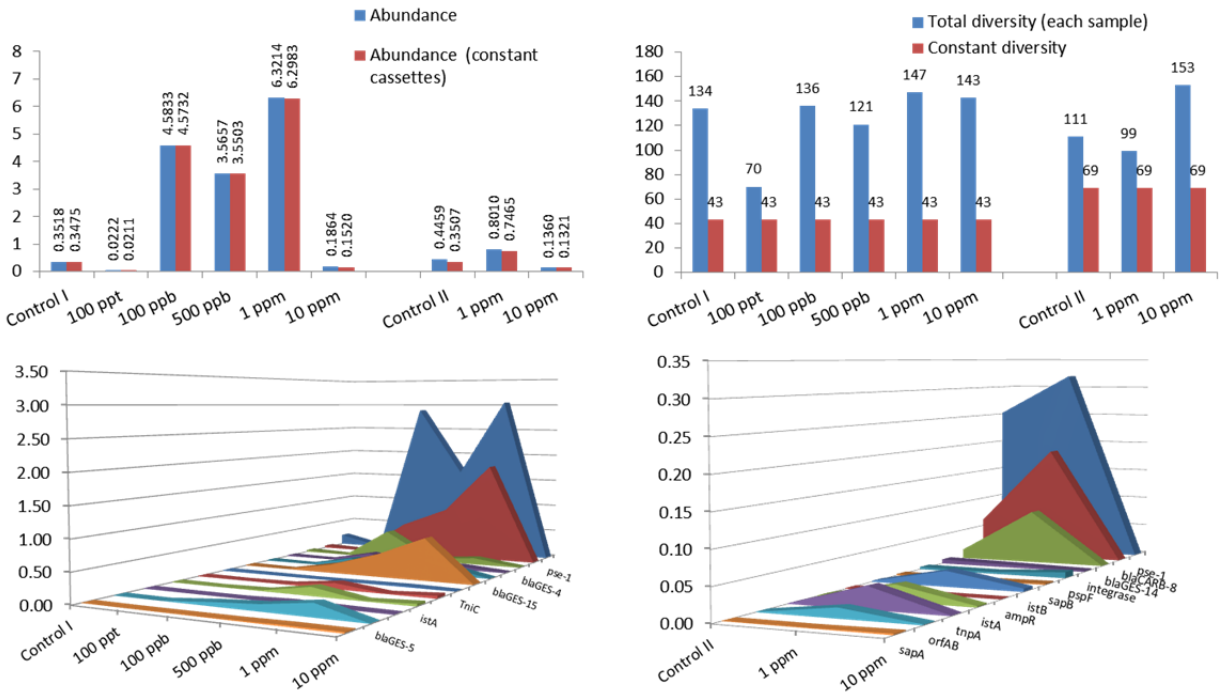


Figure 34: A - Distribution of the abundance of resistance cassette genes at Exp1 and Exp2; B - Distribution of the diversity of resistance cassettes genes at Exp1 and Exp2; C and D – Abundance of the most representative cassettes present along Exp1 and Exp2 with most of them representing the beta-lactamase genes.

In this section, the values found for the abundance and diversity of antibiotic resistance genes and genes composing integrons demonstrate a similar panorama. The increasing concentrations of triclosan promote an increasing of the abundance of cassettes genes at 1 ppm for both regimens of triclosan with the further reduction even below the detection level at 10 ppm.

Plasmids

Plasmids are common vectors of HGT and a versatile genetic tool, because of their molecular stability, and their ability, as mobile elements, able to cross the walls between cells serving to another seigneur exchanging benefits. Knowing the diversity and abundance of these elements in an environment submitted to low and high concentrations of triclosan may allow to understand their behavior upon this selective pressure.

Analyzing the control sample from both experiments at the beginning of the experiment a total of 94 plasmids were recognized at the Exp1 and 81 at Exp2. However the diversity of plasmids changed according the concentrations to which samples were submitted. At the end of each regimen Exp1 presented a total diversity of 153 plasmids and in Exp2 134. From the 94 plasmids present at the control sample from Exp1, only 40 were present in all samples (Figure 35 A) and for Exp2, only 48 were shared between all samples (Figure 35 B).

Since 40 plasmids were constant at all samples, their abundances were analyzed separately to distinguish the influence of the constant plasmids from the emerging plasmids recognized during the enrichment promoted by the samples culture. Checking this issue, the abundance of bireads at 100 ppb and 1 ppm, the most representative samples, indicated a difference of approximately 11% between the total abundance (36,17%) and the abundance of the constant plasmids (25,82%) at 100 ppb of triclosan. On the other hand, the comparison of the abundance from all plasmids and the persistent plasmids at 1 ppm did not present differences. The variations of the diversity of plasmids at Exp1 were promoted by the appearance of plasmid sequences, which were below the detection threshold at the beginning of the experiment, and the abundance of which did not disturb the overall abundance of the constant plasmids.

For Exp2, both treated samples suffered a decrease at the abundance and in the diversity of plasmids. Comparing the values of the abundance of the 48 plasmids present in the three samples of this experiment, there were no differences with respect to the abundance of bireads identified as belonging to plasmids indicating that emerging plasmids did not interfere drastically at the compositions of the samples (Figure 35 C, D).

Overall all results permit to consider that the effects caused by the presence of triclosan are dependent of the concentration of this biocide and may promote variations in the composition and abundance of genetic mobiles elements present in each sample.

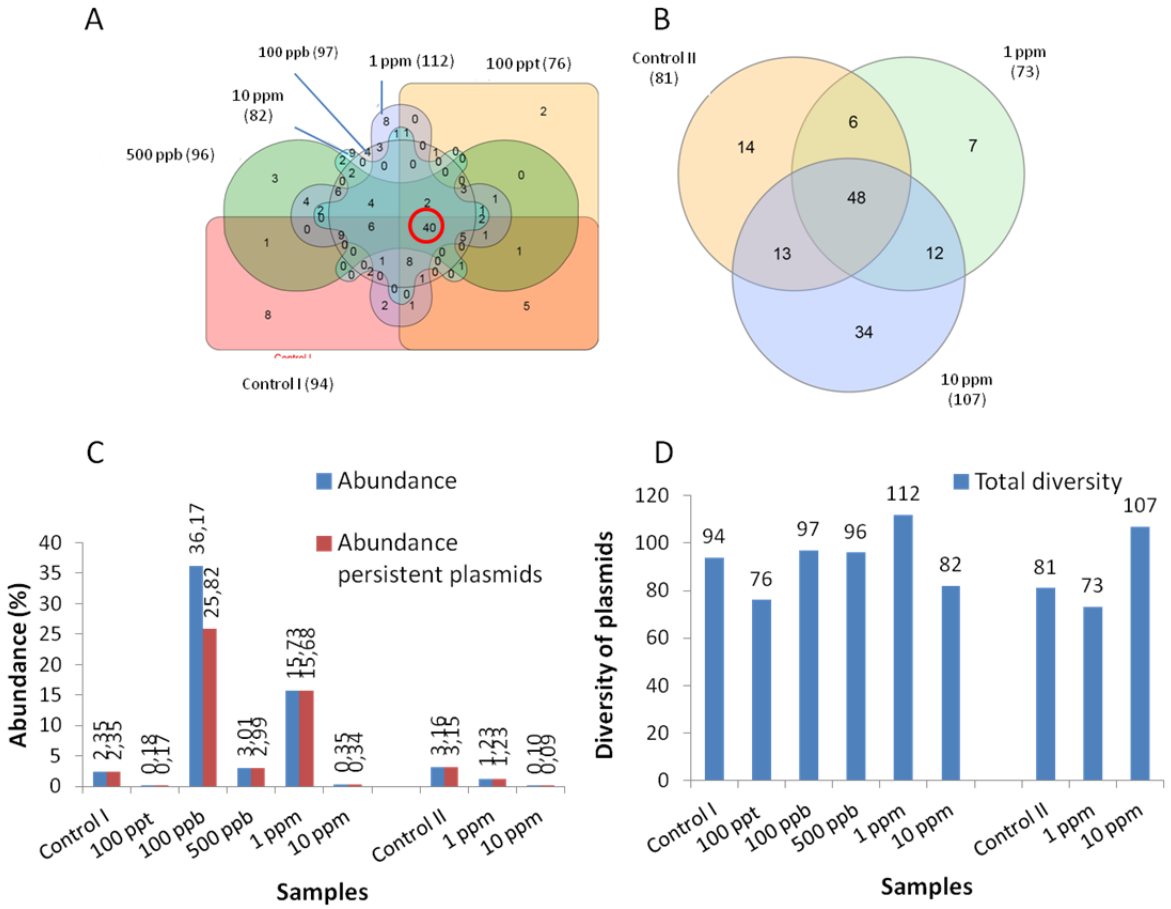


Figure 35: A - Venn diagram demonstrating the distribution of plasmids at all samples of Exp1 with the red circle indicating that 40 plasmids were recognized in all six samples; B – Venn diagram representing the distribution of plasmids present at all three samples of Exp2 recognizing 48 plasmids present in all samples of this experiment ; C – Bar chart demonstrating the abundance of bireads recognized as belonging to plasmids genomes (blue bars indicate the total abundance of bireads identified as a valid match with an e-value of 10^{-10} and red bar represent the abundance of bireads from the plasmids present in all samples of each experiment; D – Bar chart indicating the variations of diversity in each sample.

Influence of triclosan at the bacterial community structure

Shotgun sequencing of rRNA provides a new way to study communities in microbial ecology and here we used this approach to characterize the bacterial composition of nine samples obtained from the enrichment of sludge collected from a WWTP and submitted to two treatments of triclosan.

A total of 2.741.228 sequences were obtained from the sequencing facility and after filtering by quality 2.420.335 of them were used to identify the 16S rRNA genes from nine samples collected from two regimen of exposure to triclosan. For both control samples, as expected given that the waste-water treatment plant receives human residues, the typically abundant phylum was Proteobacteria followed by Actinobacteria and Bacteroidetes, respectively (Figure 36). The changes in microbial community composition were assessed in more detail by taxonomic classification of 16S sequencing reads. The abundant bacterial taxa Rhodobacterales (*Paracoccus*) and Thiotrichales (*Thiothrix*) were highly represented at the initial sample from Exp1 and Exp2, respectively and their presence were noted in all samples over the course of the experiments. In contrast, the composition of both

treatments suffered shifts for both sides, increasing the abundance in some taxa and decreasing in others.

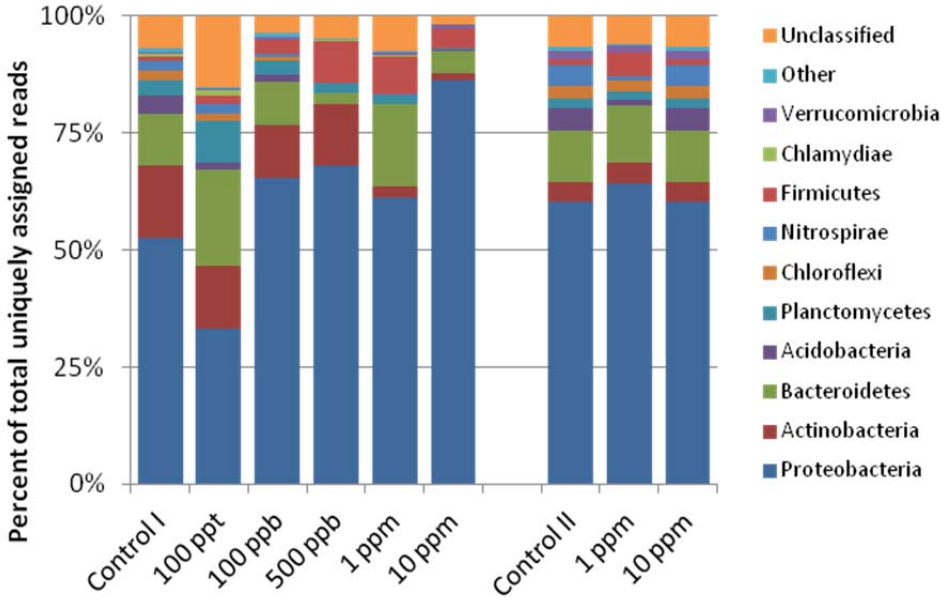


Figure 36: Microbial population composition assessed by taxonomic classification of 16S metagenomic reads at the phylum level. Only taxonomic groups representing ≥1% of abundance of total assigned reads in one dataset were included. Taxonomic groups which did not reached this abundance were grouped as "Other". The abundance for wach taxa were determined by the percentage assigned reads to each taxa divided by the total of reads assigned in each sample.

The taxa Rhodobacterales, Actinomycetales, Sphingobacteriales and Rhizobiales were represented by >30% of reads in all samples of Exp1 (Figure 37).

A decrease in abundance of Actinomycetales (15,23% in control I vs. 1,56% at 10 ppm) and Sphingobacteriales (9,13% in Control I vs. 2,59% at 10 ppm) was accompanied by the increase in the taxa Methylphilales (*Methylophilus*) accumulating approximately 41% (vs. 0,024% in Control I) in relative abundance of reads assigned, at the genus level, in the concentration of 10 ppm for Exp1.

This increase was followed by Rhizobiales (4,6% in Control I vs. 14,26% at 10 ppm) and Burkholderiales (3,20% in Control I vs. 13,25% at 10 ppm). For the sample submitted to a concentration of triclosan of 1 ppm during one week, still in Exp1, the genus *Acinetobacter* from Pseudomonadales, which harbor pathogenic organisms, was the most abundant taxon grouping 34,85% of the reads identified in this sample. In the beginning of the experiment this genus was present but in a low-level abundance (0,003%) reaching the maximum value whit the concentration of 1 ppm triclosan. Although, at 10 ppm the abundance reduced to limits near to extinction (0,04%). In the same manner, the genus *Pseudomonas* was initially low represented, increasing its abundance over the treatments.

Similar to Exp1, Exp2 presented differences at the proportion of the Phylum composition in each of the samples (Figure 36) although some populations remained as dominant with an average of 61,49% for Proteobacteria and 11,38% for Bacteroidetes for the three samples. Registered in all samples of Exp2, Rhodobacterales kept it abundance at 1 ppm (10,57% in Control II vs. 9,99% at 1 ppm). In the same manner, Sphingobacteriales

(9,61% at Control II vs. 9,35% in 1 ppm) and Actinomycetales (3,51 in Control II vs. 4,30% at 1 ppm) did not suffered drastic changes in their abundance values at low concentrations of triclosan despite the variations at the global composition of the community. In Exp2 at 1 ppm the family Methylopphilaceae reached the abundance of 26,14% of the families present in this sample, and 21,92% of the identified families at 10 ppm.

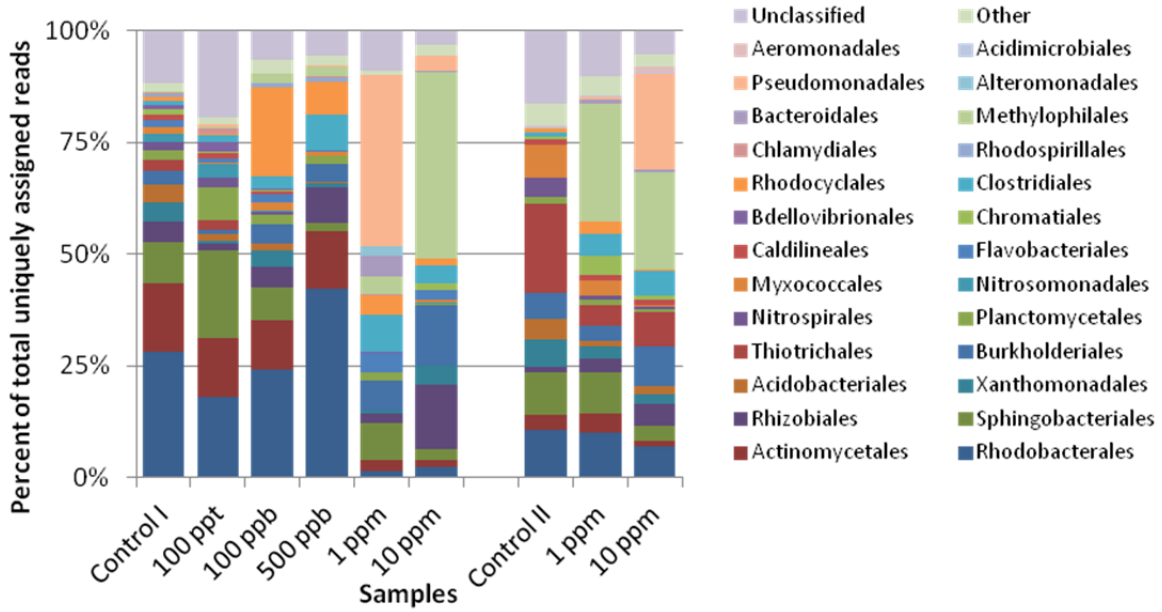


Figure 37: Microbial population composition assessed by taxonomic classification of 16S metagenomic reads at the order level. Only taxonomic groups representing ≥1% of abundance of total assigned reads in one dataset were included. Taxonomic groups which did not reached this abundance were grouped as "Other". The abundance for wach taxa were determined by the percentage assigned reads to each taxa divided by the total of reads assigned in each sample.

The sulfur-oxidizing bacteria genus *Thiothrix* and *Paracoccus* were the unique taxa which maintained their abundance over 1% in all samples of Exp2. *Thiothrix* was the most abundant taxon in the control II decreasing in the ongoing samples to 4,46% in 1 ppm and 7,5% 10 ppm. Zooming out to the family level, the composition of both treatments in Exp2 demonstrated a decrease in most taxa excepting Methylophylaceae which increased (0,12% in control II vs. 26,14% at 1 ppm) being represented by the genus *Methylophilus* (14.7%) and *Methylobacillus* (12.07%). On the other hand, Methylobacillus still increased it abundance at 10 ppm occupying the second place in abundance with 20.39% of the reads from this sample, leaded by *Pseudomonas* with 21,23% in abundance.

Discussion

"I think that in the discussion of natural problems we ought to begin not with the Scriptures, but with experiments, and demonstrations."
Galileo Galilei

5. Discussion

The increasing use of antimicrobial agents in recent years has resulted in the development of bacterial resistance to these drugs. In addition, antibiotic treatment can alter the composition of the natural microbiota and the prevalence of antimicrobial resistance (AMR) genes (Jakobsson et al., 2010; Jernberg et al., 2010). The significant clinical importance of resistance and it interest at the study of antimicrobial resistance have been investigated from different angles, analyzing the mechanisms which may contribute to the development of resistance genes in small or large scale. In a small scale we can cite the investigations of the effects caused by the presence of antimicrobials in genes, guiding the selection of resistant organisms to these compounds, as well as the dissemination of antimicrobial resistance genes present in natural bacteria (Card et al., 2015; Jernberg et al., 2010). In a large-scale overview, the effects of the antimicrobial pollutants discharged in natural environments and the effects of this substances in an entire community changing the distribution and abundance of the present organisms, eliminating part of them and increasing others. It may cause a reduction at the community diversity but, at the same time maintain the prevalent species able to survive in presence of antimicrobial compounds (Mullany, 2014).

As a chess match, where each piece has a determined role in the game, the different types of organisms, mechanisms, transmissions and evolution of antibiotic resistance exert their contribution in the whole environment to the development of antibiotic resistance. The set of genes, present in a given ecosystem and which expression can render resistance has been termed as resistome (D'Costa et al., 2006). Different approaches have been already used to investigate which pieces are the most important sustaining the structure of the ecosystem and those responsible for the maintenance of the environmental antimicrobial resistance. To this, targeted PCR, sequenced-based and functional metagenomics are the most commonly used methodologies to access the genomic information about the diversity of encoded resistance genes of bacteria which still cannot be cultured in laboratories (Moore et al., 2011; Sommer and Dantas, 2011).

The multiple layers of antibiotic resistance

It is important to mention that resistance evolution generating and improving the capabilities of one organism to survive in high concentrations of antimicrobials is not a selfish issue but rather plays the distribution of such resistance among the population. Like it happens in Public Health, what occurs for a single individual may

have consequences for the overall population. An important adjuvant for the dissemination of resistance is the interchange of genetic material between bacteria, generally involving the donation and acquisition of pieces, which creates an heterogeneous collection of elements which may display adaptive characters as resistance gene that are hierarchically linked to a complex set of transmissible factors of different types as plasmids, transposons, phages, integrons and genomic islands (Baquero, 2004). All of these elements can evolve independently, but they also form part of an hierarchical organization able to be transfer, completely or in part the elements involved, both vertically and horizontally. The multiple levels of selection this genetic structure presents, facilitate the rising of new sequences patterns, which overall tend to evolve as selective units in their host-to-host transfer (Rice, 2002; Rowe-Magnus and Mazel, 2002).

The main goal of the present work is to analyze some of the driving forces involved in evolution, spread and fixation of antibiotic resistance. For that purpose, we have explored the capability of evolution of a set of resistance genes putatively involved in quinolone resistance, we have tracked the ways that known quinolone resistance genes are integrated in plasmids (polyphyletic of monophyletic origin), whether or not the populations of bacteria that can colonize both clinical and natural ecosystems are evolving into two different branches and finally, which are the mechanisms of selection of antibiotic resistance by toxic compounds at the microbiome level. Altogether these studies enclose the different levels of genetic organization, complexity and selection with relevance in the selection of antibiotic resistant microorganisms.

On an time evolution scale, the massive explosion of antibiotic-resistant phenotypes in human and animal pathogens is a very recent event that has followed the large-scale production and use of antibiotics in clinical and veterinary medicine, agriculture, aquaculture (Aminov and Mackie, 2007; Aminov, 2009). Emergence of resistance occurred soon after the introduction of antibiotic for therapy. The most accepted hypothesis on the origin of resistance genes is that, since antibiotics are naturally-produced compounds, the producers themselves need having auto-protection mechanisms to avoid the activity of the antimicrobials they produce (Benveniste and Davies, 1973). Nevertheless, the fact that resistance genes to synthetic antibiotics as the quinolones family do exist in Nature indicates that much more potential sources of resistance genes, besides producers must been taken onto consideration.

Functional metagenomics screens allows to study this problem in depth by identifying genes by their function in an expression vector rather than by a specific

sequence used for PCR probing. Using this approach, novel antibiotic resistance genes have been identified in different environments including oral microbiota (Diaz-Torres et al., 2003), soil microbiota (Allen et al., 2009; Sommer and Dantas, 2011). Nevertheless, one question still remains in the field. Can those putative resistance genes characterized by functional metagenomics be considered as *bona fide* resistance determinants or by contrary they are just pre-resistance genes that need evolving in the treated patient to confer a fully resistant phenotype to their new bacterial host.

Chromosomal bottlenecks to the evolution of plasmid-encoded antimicrobial resistance genes

Guided by the difficulties to understand the role of antibiotic resistance in natural environments, because major studies are focused on the efficiency of antibiotics in extinct infections and pathogens, our motivation was to analyze if genes inserted in plasmids which confer low-level resistance to antibiotics may increase their potential by mutation originating higher-level resistance alleles. In this manner we have investigated along the present Thesis the effects of the second generation quinolone ciprofloxacin in selection of quinolone resistance of *E. coli* strains cells containing low-level resistance genes to this antibiotic, belonging to the pentapeptide repeat protein ¡family and obtained by means of functional metagenomic studies, from microbiomes collected from different animal farms, in order to predict the potential of such genes con confer clinically significant levels of quinolone resistance.

Nevertheless, to be fixed in a bacterial population a pre-requisite is that the expression of the gene in question does not confer a high fitness cost to the host. Most studies on fitness costs assume that in the case of plasmid-acquired resistance genes, such costs are due to the metabolic burden imposed by the replication, transcription and translation of the acquired mobile element. Nevertheless, work from our laboratory has shown these fitness costs to be allele specific. We then implemented a predictive method to analyze how stable could be the plasmids carrying the set of different *qnr*-like genes isolated from the functional metagenomic study. In all cases, the amount of cells containing the plasmids decreased along time indicating that the chances of fixation of the studied genes in the absence of antibiotic selective pressure were likely low. Actually, the decreasing in the number of carrying-plasmid cells may be justified by an intrinsic behavior of the bacterial populations, which can potentially revert to a population of bacteria susceptible to antibiotics if the antibiotic pressure is ceased. Although not studied

here, even in the presence of antibiotics, although plasmids are lost in some cases, the majority of cells may not necessarily lose their resistance properties demonstrating a reasonable structural and segregation stability of the cells (Zhang and Chisti, 1996). During culture, cells lacking the plasmid may appears coexisting and competing with the plasmid-containing population. This coexistence of plasmid-carrying and bacteria which do not carry plasmids, equally resistant to antibiotics, could be explained by the emergence of chromosomal mutations that compensate for these costs, as modifications, in our case, of the bacterial topoisomerases, which play an important role in quinolone resistance (Cambau et al., 1993; Thomas et al., 2010; Weinreich et al., 2006).

Experimental evolution in the presence of antibiotics has been largely used as a practical application for predicting evolution towards resistance to antibiotics currently in use at clinical settings, as well to new antibiotics under development. Experimental evolution has been used to track antibiotic resistance mutations but also for the prediction of the evolution of already known resistance genes in response to new antibiotics. This approach has been extremely useful for studying evolution of beta-lactamases in the present of novel beta-lactams of combinations of beta-lactams with beta-lactamases inhibitors (Novais et al., 2010; Sanders, 1989; Smith, 1970). We then decided to implement the same approach to predict the chances for evolution of quinolone resistance genes conferring low-level resistance towards high-level resistance. We found that the evolution experiment carried using quinolones provided distinct results as those when beta-lactams where used as selectors of novel beta-lactamases variants. To this class of genes is possible to establish a connection between the increasing concentrations of antibiotics and the appearance of mutations in the resistance genes However, in our case, using a increased concentrations of a second generation fluoroquinolone for selection during approximately 20 days in a set of 13 strains containing plasmid encoded quinolone resistance genes, we found that, while each of the clone acquired an increased level of quinolone resistance, the mutations responsible for this phenotype were not present in the plasmids. Indeed, cells transformed with the plasmids from the evolved cultures do not present increased level of resistance, indicating this resistance to be chromosomally encoded, likely because of mutations at the genes encoding bacterial topoisomerases.

In this regard it is important to notice that, in agreement with the information here presented, quinolone-resistant clinical isolates also presents mutations at the genes encoding bacterial topoisomerases (Bearson and Brunelle, 2015). One another aspect where *qnr* genes may be important is to favors the

emergence in mutations at chromosomal genes. Indeed, it has been earlier shown that Qnr-containing cells present a higher mutation frequency towards quinolone resistance than E. coli cells without the gene (Martínez-Martínez et al., 1998). Nevertheless, this study has been performed using a complex plasmid which contains several other genes besides *qnr*A and hence the effect on increased mutation cannot be unequivocally tracked to the presence of Qnr. If different *qnr* genes have different effect on the selection of strains presented increased resistance to quinolones, each one will present a different evolution landscape, but these landscapes should be reproducible. We did not find such conservation in the evolution along time towards resistance and hence, there is not a clear effect on the presence of different *qnr* genes on the raise of quinolone resistance in *E. coli*.

Even though we have found that high-level resistance is achieved as the consequence of chromosomal mutations, two evolved plasmids Cip7 and Cip13, slightly reduced the susceptibility of the resident cells, indicating that selection of *qnr* genes might be possible, although the consequences of this evolution is minor. Indeed, in nature plasmids carrying different *qnr* alleles of widespread genes, as *qnr*A and *qnr*B are found in many species. The existence of these alleles can be due to two different evolutionary landscapes. 1) One gene has been acquired by a given plasmid and after that the *qnr* gene has evolved under selective pressure in clinical settings. Although plasmid genes can transfer from plasmid to plasmid, the structure surrounding each of the alleles should show a certain conservation degree (monophyletic origin). 2) There have been multiple acquisition events, each one concerning one of the alleles and the structures surrounding each of the alleles are diverse (polyphyletic origin). To define the accessory structure that surround each these alleles, all plasmids, currently present at databases and containing these *qnrA* and *qnrB* alleles were analyzed. Our results demonstrated that *qnr* genes has a stable compositional structure dependent of the resistance gene type and reported as conferring low-level resistance to quinolones to their host. Each allele presented a particular arrangement in the plasmid genomes which were found, in general surrounded by transferable elements (Rice, 2002). The similarities between the alleles of *qnr* genes, and the different conformations presented by each allele type, permit to attribute a polyphyletic origin for these genes, where the evolution events occurred out of the plasmids and further incorporated to them in successive acquisition events. The small differences between each variant of *qnr* indicates a low rate of evolution for this gene and may suggest a conservative propagation when inserted in stable genomes.

We have seen that, differing to the situation with plasmids-encoded beta-lactamases, in which evolution under antibiotic treatment is a common outcome, the situation concerning Qnr determinants is not the same. There are two aspects that can justify these finding: a) the pentapeptide repeat family of proteins present a very tight structure that resembles DNA; this structural similarity to DNA is the basis of the binding of Qnr proteins to the topoisomerases and the protection of the later from the activity of the quinolones. Any change altering the structure drastically would likely impede the binding of Qnr to bacterial topoisomerases, so as impeding the selection of mutants with more effective Qnr proteins. b) If the bacterial chromosome contains genes which mutation renders resistance without an associated high fitness cost, the chances of selecting mutations in low-level resistance genes towards high level of antibiotic resistance are likely low.

Analysis of the core genome and the pangenome of *S. maltophilia*

Stenotrophomonas maltophilia, is a bacterium considered an opportunistic pathogen, which is characterized by individuals that cause infections in immunocompromised patients. Its occurrence has increased since the 90's (Fujita et al., 1996) until recent days demonstrating the importance of this opportunistic pathogen as public health problem despite *S. maltophilia* is not considered highly virulent (Gherardi et al., 2015; Härtel et al., 2012). This specie, as the great majority of bacterial species, has an environmental origin and it was described in many different habitats as soil, water bodies and associated with plants (Härtel et al., 2012; Romanenko et al., 2008; Suckstorff and Berg, 2003). The crescent frequency of *S. maltophilia* isolates exerting an important role in clinically promoted respiratory, circulatory and urinary infections indicate a wide range of adaptability of this species to occupy diverse habitats. An issue reinforced by its capability to also colonize hospital tools and utensils, as dialysis machines, catheters and prothesis had shift this bacteria to another level in the list of nosocomial pathogens.

One important aspect of *S. maltophilia* is that, besides its role as a human pathogen, it is a biotechnologically valuable species with application in confined bioreactors, but also in non-confined environmental in the field using it to protect plant species from as bio-pesticides (Pages et al., 2008), the degradation of toxic compounds as methomyl that is used as pesticide (Mohamed, 2009), the detoxification of high molecular weight polycyclic aromatic hydrocarbons as acenaphthylene, phenanthrene, chloroanilines and chlorocatechol and other compounds find in creosote-contaminated sites (Gao et al., 2013). There is then a concern on whether the use in the field of bacterial pathogens (even opportunistic

ones) may increase the risk of acquiring infections and then constitute a public health problem. Given the economical and societal consequences of forbidden (or not) the use of *S. maltophilia* for biotechnological purposes, it will be important to decipher whether environmental and clinical *S. maltophilia* constitute to separated phylogenetic branches or, by contrary, any strain can colonize a given patient, given the patient is immunocompromised or have another basal disease that favors *S. maltophilia* infection.

These situation, associated with the capability of *S. maltophilia* to resist to many antibiotics (Quinn, 1998) prompted us to perform a comprehensive study and characterization of 24 strains of *S. maltophilia* with different origins, sequencing for the first time 21 isolates and establishing their core and pangenome, as well the number of accessory genes presented by all strains sequenced in this Thesis together with three closed genomes available at the time this work began.

In a first insight, the complete genome of the strain *S. maltophilia* D457 revealed a similar composition and genomic structure conservation when compared with the other three available complete genomes of this bacterial species. On the other hand, its genome presented exclusive Genomic Islands (GEIs). Despite all the GEIs are present in the D457 strain, not all genes present in each of the GEIs were exclusively present in this strain. Genes present in some GEIs demonstrated to be shared with other strains but they were not included at the core genome, since they were not shared by all of the isolates and have then be recognized as part of *S. maltophilia* accessory genome.

When looking on more detail the core genome of clinical and environmental strains with closed genomes, the number of genes shared by all strains (2742 genes) presented a percentage of 61-65% of the total genes harbored by the four strains. This value is much lower as compared with other free-living bacteria, as *Pseudomonas aeruginosa*, which present 5316 coding sequences in the core genome representing approximately 93% of the total genes of this specie (Ozer et al., 2014). At the other side of the core genome, we can find the multi-specialist with *E.coli* with 1472 core genes from a total of 4498 genes representing 32% of the genes for this specie (Rasko et al., 2008). The fact that the core genome if *S. maltophilia* is lower than other free-living bacteria may indicate the existence, as in *E. coli*, of different lineages, each one adapted to colonize a given ecosystem, and that have evolved through the acquisition of adaptive genes by means of different HGT events.

When the 20 draft genomes of the new sequenced isolates of *S. maltophilia* were included in the analysis, the results obtained demonstrate an increasing size of the pangenome (from 5991 to 12214 genes) for this specie, accompanied by the

decreasing in the number of genes shared by all strains present in the study. The reduction of genes shared in the core genome, from 2742 to 1875, suggesting that *S. maltophilia* presents a small core genome, corresponding in range, to 42% for the 24 strains analyzed. Nevertheless, this reduction in number of genes belonging to the core genome could be explained (at least in part) by the fragmentation of the 20 new sequenced genomes. To minimize the effects of fragmentation of the genome and the high number of contigs that could discharge interrupted genes, the softcore were calculated considering the genes shared by at least 90% of the isolates, increasing this number to 2585. This value is closer to the previous value obtained using the four reference strains to establish the core genome of *S. maltophilia* based on complete genomes. In this manner, we highlight the importance to take in count the softcore genome when working with a large set of isolates which present draft genomes.

Evaluating the genomic distribution of the elements that form the pangenome from clinical and environmental strains, the 22 clinical isolates presented a total of 9542 clusters versus the 8747 formed by the environmental strains. It indicates that the clinical strains of *S. maltophilia* evaluated in this study are slightly more diverse in genes than the other group. Analyzing separately the components of the pangenome, the core genome did not presented a remarkably difference in number, sustaining the idea of a reasonable estimation of the core genome will be approximately 49% of the total genes for the clinical and environmental groups. The contribution of the softcore and shell genome comprised, for those groups, were very similar with 50% to the environmental and 47% for the clinical strains. When observing the cloud genome of each group the was possible to identify and understand the origin of the difference between the number of clusters formed by the clinical and environmental strains. This difference may suggest the presence of specific genes in clinical isolates related to the adaptation to adverse conditions presented in clinical settings, as antibiotic pressure.

Clinical and environmental strains of *S. maltophilia* do not form different phylogenetic branches

It was confirmed scanning strategic genes responsible for virulence, motility and infection described previously in other works (Fouhy et al., 2007; Pol Huedo et al., 2014), and once more it was not possible to distinguish a clear separation between the groups with environmental and clinical strains found together sharing the same clusters. being confirmed further using a phylogenetic approach

comparing the CDS composition of each strain and any specific origin-branch group were identified.

When using the CDS to establish the similarities between the analyzed strains, it demonstrate that clinical and environmental strains were separated in clusters containing in all of them, both class of origin strains. Despite the strains mix themselves, curiously the complete genomes of *S. maltophilia* D457, K279a, R551-3 and JV3 were grouped together. It may be consistent with the fact that these genomes are closed and despite their differences any gene was interrupted in its genome. At this same branch the clinical strain *S. maltophilia* D457 was closely related to the environmental isolates JV3 and R551-3 enabling to those the putative status of "infectious strains" based on the CDS composition proximity. Although, further analyzes must be considered to confirm this plasticity in the strains. Curiously, the three strains obtained from sewages were grouped together with a clinical strain collected from a blood sample.

The try to give an overview on a broad spectrum of *S. maltophilia* isolates from different sources characterizing all 24 strains with respect to a set of strategic genes present in the genome, some of them related with the motility of this organism and mechanisms responsible for the transport and secretion of transport-effectors proteins through the membrane did not indicated a trustful way to separate the isolates by origin. Using a presence/absent matrix to calculate the similarities of these specific strategic genes and further compare the composition of cited strains, the clusters formed did not represented an aggrupation excepting the similarities of the previous organization saw at Figure 27. Using this genes as criteria to classify the strains, a major group formed using a threshold of 50% of similarity were constituted by both clinical and environmental isolates. This findings, associated with the fact that clinical strains are submitted to a constant pressure of antibiotics it will be the unique factor which promote the major differences but concerning with the concentrations of antibiotic accepted by resistant strains.

Complementing the approaches used before, when analyzing the MICs of the isolates with 12 antibiotics, the results obtained indicated a slight separation of environmental and clinical strains when the isolates were compared between then and depending on the antibiotics included in the analysis. In this manner, the environmental strains had their MICs centered in a small range of concentration and in some cases lower with respect to the same antibiotic applied to clinical strains. This behavior could be explained by the exhaustive selective pressure under antibiotics suffered by the clinical strains and not an exclusive property inherent from the strains which belong to this groups. As already known, alleles variants

present in the genome can affect the resistance mechanisms without interfere at the CDS composition and no correlations could be made to separate the groups with this approach.

Concerning to the mechanisms which may interfere at the virulence of some bacteria, for example the Diffusible Signal Factor Quorum Sensing (DSF-QS) gene *rpfF* responsible for the regulation of DSDF-AQ, we characterized 24 S. maltophilia strains isolated from diverse sources. The characterization based on the rpfF gene revealed the presence of two variants in all strains analyzed differing of the analysis performed by previous work which described the lack of *rpfF* gene in some strains of *S. maltophilia* (Pompilio et al., 2011). Analyzing our results, the proportion of 50% to both groups carrying the *rpfF* variants differs from the study carried using 171 strains , which four of them were present in our analysis (D457, K279a, R551-3 and JV3). In this study, Huedo and friends found the proportion of 60 and 40% for the variants rpfF1 and rpfF2, respectively (P. Huedo et al., 2014).

These results suggests that some subgroups from certain ecological and clinical origins may exhibit similarities, as well peculiar differences, concerning to their relevance in each type of environment. In this case, specific genes could be present in the genome of these strains to support the occupation and colonization in different substrates and conditions but they are not restrict to a determined context.

Antimicrobial resistance genes are equally distributed in clinical and environmental *S. maltophilia* isolates

We have studied the behavior of clinical and environmental isolates of *S. maltophilia* with respect to the susceptibility to common antibiotics used in clinical treatments. Our results demonstrated that both lineages present a wide range of resistance depending on the antibiotic applied. In general, the overall MIC values obtained for clinical were higher as compared with the values obtained for the environmental group. Despite the clinical isolates present higher MICs than environmental strains, this last group presented specific isolates, which have MIC values similar to the clinical resistance set. These exceptions increased the range of resistance for the environmental strains and at the same time, together with the results shown above, may support the idea that environmental strains are potential opportunistic resistant pathogens. Normalizing the MICs results obtained with the MIC50 for all strains, it was possible to note differences and internal separation of groups, but using the external clinical strain S. maltophilia D457 for the normalization, the distribution of MICs resulted in a mix of the different isolates,

from both clinical and experiment origin and with no clear separation between groups. These small differences between the MICs for clinical and environmental groups support that *S. maltophilia* has intrinsic antimicrobial resistance mechanisms independently of their origin. In addition, the increased MIC values found for clinical strains (and some environmental ones also) may be the consequence of living under the constant pressure of antibiotics used in clinical treatments and not promoted by the variations at the number of resistance genes because both groups did not presented differences in the number of such genes.

Supported by the results and similarity between the strains, the proliferation of the environmental strains in the clinical ambient may act as opportunistic infectious agents. A major expectation from repeated antibiotic exposure of the microbiota is an increasing abundance of antibiotic resistance determinants, caused by the growth advantage of resistant organisms during antibiotic treatment. If it really occurs efficiently, the environmental strains may play an important role in the infection of compromised patients.

The concentration of triclosan, and not just its presence, may guide the selection of resistant organisms in the environment

In many developing countries the unregulated sale and dispensing of antimicrobial compounds is very common. Thus, it is important not only to consider the contribution of hospital effluents at the increasing antibiotic resistance promoted by the discharge of antimicrobial compounds, but also the contribution of the general community to the input of antibiotic-resistant bacteria to the aquatic environment. In addition to this situation, other compounds with antimicrobial activity (the biocides) are widely used for several purposes without any restriction in any country. One of the most widely used biocide is the triclosan. One worrying consequence of its use is that it might contribute to reduce the susceptibility to clinically antimicrobials, due to either cross-resistance or co-resistance mechanisms (Chuanchuen et al., 2003; Yazdankhah et al., 2006). Largely used in household products and in contrast to other biocides, triclosan at low concentrations acts similarly to antibiotics on the cellular target enoyl-acyl carrier protein reductase (FabI), an essential enzyme in bacterial fatty acid synthesis (Fan et al., 2002; Schweizer, 2001).

The frequent presence of triclosan in anthropic environments and the possibility to be an inducer of antimicrobial resistance in environmental settings guided our curiosity to understand the effects of this compound in a controlled natural environment. To this goal, culture of samples of activated sludge collected

from WWTP were sequenced and analyzed with respect to the composition and diversity of transferable genetic elements carrying resistance genes, as well the taxonomic composition of the samples along the experiment.

If triclosan selects antibiotic resistant microorganisms, a monotonic increase in the number of resistance genes as a function of triclosan concentration would be expected; the highest triclosan concentration would render also the highest amount of antibiotic genes. What we observe however is not such simple outcome.

Observing the results obtained from the analysis of the sequencing shown, at Exp1, a drastic increase in abundance for antimicrobial resistance genes in low concentrations of triclosan. An abundance of 15% was reached at 100 ppb, the same concentration where the genus *Methyloversatilis*, common used in bioremediation of such hazardous pollutants, trichloroethylene and other halogenated methane and in nitrogen removal from sewage and fertilized soils (Kalyuzhnaya et al., 2006), had its largest abundance (see below). It is important to recall here that this increase in mainly for plasmid-encoded genes; the differences in abundance and diversity in diversity of genes might depend of the number of copies in each cell (in other words the type of plasmid, low or high copy number), but it can indicate the enrichment of a small group of carrying-plasmid cells able to grow in this concentration of triclosan. The fact that with the concentration of 500 ppb the abundance of resistance genes returned to it "normal" level, indicates that enrichment of plasmid-encoded resistance genes is highly dependent on triclosan concentrations. Mild concentrations of this biocide select for antibiotic resistance, whereas high concentrations do not have selective value. Guided by the abundance of beta-lactamase type TEM genes present in all samples, a screening to reconstruct these genes was performed using the raw sequences to know which allele of this gene were present in the samples. Applying this strategy, it was possible to identify 20, from the 53 previously identified, complete genes of beta-lactamase genes at 100 ppb of triclosan. It could indicate that the presence of triclosan in low concentrations facilitates the selection of bacteria carrying beta-lactamase genes which demonstrate a link between the presence of biocides and the putative resistance to antibiotics considering that these present genes are functional (Levy, 2002; Russell, 2000). Opposite ti this situation, in absence of the initial enrichment step promoted by the successive passes, mainly at 100 ppb, Exp2 experimented a reduction of the initial abundance and diversity, as well the TEM beta-lactamase genes in both treatments.

Variations found at the distribution and abundance of cassettes genes during the treatments of Exp1 and Exp2 suggested an increase in diversity of these

elements. However, this variations in number did not represent a significant change in the abundance of these genes. Listing the cassettes shared by all samples separated by experiments, it was possible to recognize that the major abundance distribution was carried by the shared genes. The appearance of new cassettes distinct from the shared for all samples of each experiment did not influence the total abundance. It could indicate that the in a closed system, the prevalent cassettes will be maintained despite the emergence of "new" organisms, which prevalence was before below the limits of detection by using our metagenomic approach.

Once recognized the genes present in the samples and the resistance cassette genes, it was necessary to correlate these data with the plasmid diversity and abundance found in each sample. Observing the behavior of this mobile element, we found an increasing in abundance at 100 ppb for Exp1. This is consistent with the finding of a higher abundance of resistance genes at this specific concentration of the biocide. For Exp2, both treatments suffered a reduction at the abundance level suggesting that the absence of previous steps were critical to reduce the abundance of species carrying plasmids. In addition, the concentration used in this second experiment were above the selective concentration of triclosan (100 ppb) already found in Exp1. Otherwise, it indicates the most of carrying-plasmids cells were not able to survive under high concentrations of triclosan remaining only those presenting high intrinsic to triclosan, which likely do not belong to the major groups carrying antibiotic resistance genes. Distinct to Exp2, the Exp1 presented two stages of plasmid distribution with marked increasing in abundance at 100 ppb and 1 ppm of triclosan. The increasing at the total abundance of plasmids overcame the abundance of persistent plasmids which might favor the selection of specific carrying-plasmid bacteria with it further decrease into non-detectable level.

Bacterial communities are influenced when exposed to triclosan: The clonal selection window.

The presence of triclosan in different soluble products and it low solubility in water turn this biocide problematic compound when it is drained from homes, industries, hospitals and farms to WWTPs. The constant input of this biocide released by urban centers may affect the stability of microbial communities selecting organisms resistant to triclosan. Some authors suggest the influence of triclosan and other biocides as an inductor of antimicrobial resistance (Russell, 2000; Schweizer, 2001; Yazdankhah et al., 2006). This study highlight the differential

consequences that the presence of low or high concentrations of triclosan exert on the sludge microbiota.

Some years ago, the concept of antibiotic selection window was proposed to describe the range of concentrations at which a given antibiotic exert a selective pressure over a given bacteria. These range goes from the minimal inhibitory concentration towards the antibiotic concentration that inhibits the growth of a single-step antibiotic resistant mutant. Although some recent works indicate that the selection window can expand to lower concentrations of antibiotics, this concept still remains. Bases in this concept, the thesis introduce the concept of clone selection window, understood as the range of concentrations of any given antimicrobial, in which a specific clone (or species) present in a microbiota is enriched (selected) over the others. In the case that the selected clone(s) harbor plasmids containing antibiotic resistance genes, this will cause the selection of resistance by a mechanism that is neither co-resistance not cross-resistance, the only mechanism of selection of antibiotic resistance by biocides so far described.

Variations at the composition of the samples studied in this work indicated that indeed, the presence of triclosan in a given environment influenced the composition and abundance of taxa, but that changes are specific for each triclosan concentration. Both experiments described in this Thesis sustained the idea that the influence exerted by the presence of triclosan in the microbiota of sludge collected from WWTPs do not drive monotonically the increase in antibiotic resistance. In general lines, the variations at the diversity of organisms detected in the samples may indicate that the concentrations applied in specific selective concentration (clonal selection window), may just be selecting those microorganisms that frequently carry mobile elements containing antimicrobial resistance genes.

Conclusions

"Life is the art of drawing sufficient conclusions from insufficient premises."
Samuel Butler

6. Conclusions

1. The analyzed plasmid-encoded quinolone resistance genes confer a fitness costs, which leads to the loss of the plasmids in the absence of selection pressure.
2. Non-chromosomal genes which confer low-level resistance to quinolones do not evolve towards high-level resistance because chromosomal mutations allow to acquire high-level resistance.
3. The analysis of the genomic composition of 24 strains of *S. maltophilia* suggests that this species has an open pangenome able to continuously incorporate new genes.
4. The analysis of 24 genomes of *S. maltophilia* from different origins show that environmental and clinical isolates of this species do not form two phylogenetic groups, neither when analyzed with respect to the CDS content of each isolate, nor when observing the groups formed after the identification of strategic genes and the mechanisms of cell communication.
5. Clinical isolates of *S. maltophilia* present an overall higher resistance to antibiotics. Nevertheless, environmental and clinical isolates do not clearly split in two groups and the number of resistance genes in each of the categories is similar, which suggest that resistance has been acquired after treatment and is not a pre-requisite for an infective behavior.
6. The presence of triclosan into a controlled culture of sludge samples collected from a Waste Water Treatment Plant revealed an increasing in abundance of Genetic Mobile Elements (GME) and resistance genes at low concentrations (100 and 500 ppb), despite the reduction in diversity indicating the prevalence of specific taxa at these concentrations.
7. The exposure to triclosan alters the taxonomic composition of the microbiota of waste water treatment plants.
8. The effect of triclosan in selecting antibiotic resistance genes is likely due to the selection of the microorganisms containing such genes at specific biocide concentrations.
9. The term "clonal selection window" is introduced as the range of concentrations of an antimicrobial that selects or maintain specific species or clones in a given microbiota.
10. The use of draft genomes to calculate the pangenome and core genome of a large set of genomes must take into account the fact that fragmented sequences will mask the genes which could be interrupted between two fragmented contigs underestimating the real number of shared genes between all the isolates present in the analysis.

7. Conclusiones

1. Los genes de resistencia a quinolonas analizados presentes en los plásmidos confieren un coste fisiológico (*fitness cost*), lo que conduce a la expulsión del plásmido en ausencia de presión selectiva.
2. Los genes no cromosómicos que confieren resistencia de bajo nivel a las quinolonas no evolucionan hacia la resistencia de alto nivel porque las mutaciones cromosómicas permiten adquirir dicha resistencia.
3. El análisis de la composición genómica de 24 cepas de *S. maltophilia* sugiere que esta especie tiene un pangenome abierto capaz de incorporar continuamente nuevos genes.
4. El análisis de los 24 genomas de S. maltophilia procedentes de diferentes orígenes demostró que las cepas ambientales y clínicas de esta especie no forman dos grupos filogenéticos distintos cuando se analizan con respecto al contenido de las CDS de cada aislado y tampoco al observar los grupos formados después de la identificación de ciertos genes estratégicos y de los mecanismos de comunicación celular.
5. Los aislados clínicos de *S. maltophilia* presentaron, en general, una mayor resistencia a los antibióticos. Sin embargo, los aislados ambientales y clínicos no se separaron claramente en dos grupos y el número de genes de resistencia en cada una de las categorías es similar, lo que sugiere que la resistencia ha sido adquirida después del tratamiento y no es un requisito previo para el éxito infectivo de las cepas.
6. La presencia de triclosan en un cultivo controlado de muestras de lodos recogidos de una Planta de Tratamiento de Aguas Residuales reveló un aumento en la abundancia de Elementos Genéticos Móviles (EGM) y genes de resistencia a bajas concentraciones (100 y 500 ppb) a pesar de la reducción de la diversidad, indicando la prevalencia de taxones específicos en estas concentraciones.
7. La exposición al triclosan altera la composición taxonómica de la microbiota de las Estaciones de Tratamiento de Aguas Residuales (ETAR).
8. El efecto de triclosan en la selección de los genes de resistencia a antibióticos se debe posiblemente a la selección de los microorganismos que contienen los genes de resistencia examinados en concentraciones específicas de biocidas.
9. El término "ventana de selección clonal" se presenta como el rango de concentraciones de un agente antimicrobiano que selecciona algunas especies o clones específicos en una microbiota dada.
10. El uso de genomas no cerrados para calcular el pangenoma y genoma core de un conjunto de cepas debe tener en cuenta las secuencias fragmentadas que enmascarar ciertos genes que podrían estar interrumpidos entre dos contigs, subestimando el número real de genes compartidos entre todos los aislados presentes en las muestras analizadas.

Bibliography

8. Bibliography

Adamek, M., Linke, B., Schwartz, T., 2014. Virulence genes in clinical and environmental Stenotrophomas maltophilia isolates: a genome sequencing and gene expression approach. Microb. Pathog. 67-68, 20–30. doi:10.1016/j.micpath.2014.02.001

Adamek, M., Overhage, J., Bathe, S., Winter, J., Fischer, R., Schwartz, T., 2011. Genotyping of environmental and clinical Stenotrophomonas maltophilia isolates and their pathogenic potential. PLoS One 6, e27615. doi:10.1371/journal.pone.0027615

Ahmed, N., Dobrindt, U., Hacker, J., Hasnain, S.E., 2008. Genomic fluidity and pathogenic bacteria: applications in diagnostics, epidemiology and intervention. Nat. Rev. Microbiol. 6, 387–394. doi:10.1038/nrmicro1889

Aiello, A.E., Marshall, B., Levy, S.B., Della-Latta, P., Lin, S.X., Larson, E., 2005. Antibacterial cleaning products and drug resistance. Emerg. Infect. Dis. 11, 1565–70. doi:10.3201/eid1110.041276

Alcaraz, L.D., 2014. Pan-genomics: Unmasking the gene diversity hidden in the bacteria species. Front. Ecol. Evol. Complex. 1–7. doi:10.7287/peerj.preprints.113v1

Alikhan, N.-F., Petty, N.K., Ben Zakour, N.L., Beatson, S. a, 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics 12, 402. doi:10.1186/1471-2164-12-402

Allen, H.K., 2014. Antibiotic resistance gene discovery in food-producing animals. Curr. Opin. Microbiol. 19C, 25–29. doi:10.1016/j.mib.2014.06.001

Allen, H.K., Donato, J., Wang, H.H., Cloud-Hansen, K. a, Davies, J., Handelsman, J., 2010. Call of the wild: antibiotic resistance genes in natural environments. Nat. Rev. Microbiol. 8, 251–9. doi:10.1038/nrmicro2312

Allen, H.K., Moe, L. a, Rodbumrer, J., Gaarder, A., Handelsman, J., 2009. Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. ISME J. 3, 243–51. doi:10.1038/ismej.2008.86

Alonso, A., Martinez, J.L., 2001. Expression of Multidrug Efflux Pump SmeDEF by Clinical Isolates of Stenotrophomonas maltophilia. Antimicrob. Agents Chemother. 45, 1879–1881. doi:10.1128/AAC.45.6.1879-1881.2001

Alonso, A., Martínez, J.L., 2000. Cloning and characterization of SmeDEF, a novel multidrug efflux pump from Stenotrophomonas maltophilia. Antimicrob. Agents Chemother. 44, 3079.

Alonso, A., Sánchez, P., Martínez, J.L., 2001. Environmental selection of antibiotic resistance genes. Environ. Microbiol. 3, 1–9. doi:10.1046/j.1462-2920.2001.00161.x

Aminov, R.I., 2009. The role of antibiotics and antibiotic resistance in nature. Environ. Microbiol. 11, 2970–2988. doi:10.1111/j.1462-2920.2009.01972.x

Aminov, R.I., Mackie, R.I., 2007. Evolution and ecology of antibiotic resistance genes. FEMS Microbiol. Lett. 271, 147–161. doi:10.1111/j.1574-6968.2007.00757.x

Amos, G.C. a., Zhang, L., Hawkey, P.M., Gaze, W.H., Wellington, E.M., 2014. Functional metagenomic analysis reveals rivers are a reservoir for diverse antibiotic resistance genes. Vet. Microbiol. doi:10.1016/j.vetmic.2014.02.017

Andersson, D.I., Hughes, D., 2011. Persistence of antibiotic resistance in bacterial populations. FEMS Microbiol. Rev. 35, 901–911. doi:10.1111/j.1574-6976.2011.00289.x

Angulo, F.J., Baker, N.L., Olsen, S.J., Anderson, A., Barrett, T.J., 2004. Antimicrobial use in agriculture: controlling the transfer of antimicrobial resistance to humans. Semin. Pediatr. Infect. Dis. 15, 78–85. doi:10.1053/j.spid.2004.01.010

Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M.J., Natale, D. a, O'Donovan, C., Redaschi, N., Yeh, L.-S.L., 2004. UniProt: the Universal Protein knowledgebase. Nucleic Acids Res. 32, D115–D119. doi:10.1093/nar/gkh131

Aziz, R.K., Bartels, D., Best, A. a, DeJongh, M., Disz, T., Edwards, R. a, Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R. a, McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O., 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9, 75. doi:10.1186/1471-2164-9-75

Bader, L., Riedel, K.G., Maydl, G., Ritter, E., Wirsing von König, C., Meroe, A., Billing, J., Hensel, G., 1999. Postoperative Stenotrophomonas maltophilia. Augeninfektionen durch beim Hersteller Kontam. intraokulare Abstract, 241–245.

Baker-Austin, C., Wright, M.S., Stepanauskas, R., McArthur, J. V, 2006. Co-selection of antibiotic and metal resistance. Trends Microbiol. 14, 176–82. doi:10.1016/j.tim.2006.02.006

Baquero, F., 2004. From pieces to patterns: evolutionary engineering in bacterial pathogens. Nat. Rev. Microbiol. 2, 510–518. doi:10.1038/nrmicro909

Baquirin, M.H.C., Barlow, M., 2008. Evolution and recombination of the plasmidic qnr alleles. J. Mol. Evol. 67, 103–110. doi:10.1007/s00239-008-9131-3

Beaber, J.W., Hochhut, B., Waldor, M.K., 2004. SOS response promotes horizontal dissemination of antibiotic resistance genes. Nature 427, 72–74. doi:10.1038/nature02241

Bearson, B.L., Brunelle, B.W., 2015. Fluoroquinolone induction of phage-mediated gene transfer in multidrug-resistant Salmonella. Int. J. Antimicrob. Agents 46, 201–204. doi:10.1016/j.ijantimicag.2015.04.008

Bengtsson-Palme, J., Boulund, F., Fick, J., Kristiansson, E., Larsson, D.G.J., 2014. Shotgun metagenomics reveals a wide array of antibiotic resistance genes and mobile elements in a polluted lake in India. Front. Microbiol. 5, 1–14. doi:10.3389/fmicb.2014.00648

Benveniste, R., Davies, J., 1973. Aminoglycoside antibiotic-inactivating enzymes in actinomycetes similar to those present in clinical isolates of antibiotic-resistant bacteria. Proc. Natl. Acad. Sci. U. S. A. 70, 2276–2280. doi:10.1073/pnas.70.8.2276

Berg, G., Minkwitz, A., Berg, G., 2001. Comparison of antifungal activities and 16S ribosomal DNA sequences of clinical and environmental isolates of Stenotrophomonas maltophilia. J. Clin. Microbiol. 39, 139–145. doi:10.1128/JCM.39.1.139

Berg, G., Roskot, N., Smalla, K., 1999. Genotypic and phenotypic relationships between clinical and environmental isolates of Stenotrophomonas maltophilia. J. Clin. Microbiol. 37, 3594–600.

Berlemont, R., Delsaute, M., Pipers, D., D'Amico, S., Feller, G., Galleni, M., Power, P., 2009. Insights into bacterial cellulose biosynthesis by functional metagenomics on Antarctic soil samples. ISME J. 3, 1070–1081. doi:10.1038/ismej.2009.48

Berman, H.F., Riley, L.W., 2013. Identification of novel antimicrobial resistance genes from microbiota on retail spinach. BMC Microbiol. 13, 272. doi:10.1186/1471-2180-13-272

Bester, K., 2005. Fate of triclosan and triclosan-methyl in sewage treatment plants and surface waters. Arch. Environ. Contam. Toxicol. 49, 9–17. doi:10.1007/s00244-004-0155-4

Braoudaki, M., Hilton, A.C., 2004a. Low level of cross-resistance between triclosan and antibiotics in Escherichia coli K-12 and E. coli O55 compared to E. coli O157. FEMS Microbiol. Lett. 235, 305–309. doi:10.1016/j.femsle.2004.04.049

Braoudaki, M., Hilton, A.C., 2004b. Adaptive Resistance to Biocides in Salmonella enterica and Escherichia coli O157 and Cross-Resistance to Antimicrobial Agents. J. Clin. Microbiol. 42, 73–78. doi:10.1128/JCM.42.1.73-78.2004

Brooke, J.S., 2012. Stenotrophomonas maltophilia: an emerging global opportunistic pathogen. Clin. Microbiol. Rev. 25, 2–41. doi:10.1128/CMR.00019-11

Brown, S.A., 1996. Fluoroquinolones in animal health. J. Vet. Pharmacol. Ther. 19, 1–14. doi:10.1111/j.1365-2885.1996.tb00001.x

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421. doi:10.1186/1471-2105-10-421

Cambau, E., Bordon, F., Collatz, E., 1993. Novel gyrA point mutation in a strain of Escherichia coli resistant to fluoroquinolones but not to nalidixic acid . Novel gyrA Point Mutation in a Strain of Escherichia coli Resistant to Fluoroquinolones but Not to Nalidixic Acid 37, 1247–1252. doi:10.1128/AAC.37.6.1247.Updated

Cantwell, M.G., Wilson, B. a., Zhu, J., Wallace, G.T., King, J.W., Olsen, C.R., Burgess, R.M., Smith, J.P., 2010. Temporal trends of triclosan contamination in dated sediment cores from four urbanized estuaries: Evidence of preservation and accumulation. Chemosphere 78, 347–352. doi:10.1016/j.chemosphere.2009.11.021

Card, R.M., Mafura, M., Hunt, T., Kirchner, M., Weile, J., Rashid, M.-U., Weintraub, A., Nord, C.E., Anjum, M.F., 2015. The impact of ciprofloxacin and clindamycin administration on Gram-negative bacteria isolated from healthy volunteers and characterizing the resistance genes they harbour. Antimicrob. Agents Chemother. AAC.00068–15. doi:10.1128/AAC.00068-15

Cascales, E., Cambillau, C., 2012. Structural biology of type VI secretion systems. Philos. Trans. R. Soc. B Biol. Sci. 367, 1102–1111. doi:10.1098/rstb.2011.0209
Cash, R. a., Narasimhan, V., 2000. Impediments to global surveillance of infectious diseases: Consequences of open reporting in a global economy. Bull. World Health Organ. 78, 1358–1367. doi:10.1057/palgrave.development.1110097
Chevreux, B., Wetter, T., Suhai, S., 1999. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information, in: Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB). pp. 45–56.
Choi, Y., Chan, a. P., 2015. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics 1–3. doi:10.1093/bioinformatics/btv195
Chuanchuen, R., Karkhoff-Schweizer, R.R., Schweizer, H.P., 2003. High-level triclosan resistance in Pseudomonas aeruginosa is solely a result of efflux. Am. J. Infect. Control 31, 124–127. doi:10.1067/mic.2003.11
Cirz, R.T., Chin, J.K., Andes, D.R., De Crécy-Lagard, V., Craig, W. a., Romesberg, F.E., 2005. Inhibition of mutation and combating the evolution of antibiotic resistance. PLoS Biol. 3, 1024–1033. doi:10.1371/journal.pbio.0030176
Clardy, J., Walsh, C., 2004. Lessons from natural molecules. Nature 432, 829–837. doi:10.1038/nature03194
Contreras-Moreira, B., Vinuesa, P., 2013. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. Appl. Environ. Microbiol. 79, 7696–7701. doi:10.1128/AEM.02411-13
Criscuolo, A., Brisse, S., 2013. AlienTrimmer: A tool to quickly and accurately trim off multiple short contaminant sequences from high-throughput sequencing reads. Genomics 102, 500–506. doi:10.1016/j.ygeno.2013.07.011
Crossman, L.C., Gould, V.C., Dow, J.M., Vernikos, G.S., Okazaki, A., Sebaihia, M., Saunders, D., Arrowsmith, C., Carver, T., Peters, N., Adlem, E., Kerhornou, A., Lord, A., Murphy, L., Seeger, K., Squares, R., Rutter, S., Quail, M.A., Rajandream, M.-A., Harris, D., Churcher, C., Bentley, S.D., Parkhill, J., Thomson, N.R., Avison, M.B., 2008. The complete genome, comparative and functional analysis of Stenotrophomonas maltophilia reveals an organism heavily shielded by drug resistance determinants. Genome Biol. 9, R74. doi:10.1186/gb-2008-9-4-r74
Cunningham, V.L., Buzby, M., Hutchinson, T., Mastrocco, F., Parke, N., Roden, N., 2006. Effects of Human Pharmaceuticals on Aquatic Life: Next Steps. Environ. Sci. Technol. 40, 3456–3462. doi:10.1021/es063017b
D'Costa, V.M., McGrann, K.M., Hughes, D.W., Wright, G.D., 2006. Sampling the antibiotic resistome. Science 311, 374–377. doi:10.1126/science.1120800
Davies, J., Davies, D., 2010. Origins and evolution of antibiotic resistance. Microbiol. Mol. Biol. Rev. 74, 417–33. doi:10.1128/MMBR.00016-10
Demain, A.L., 2009. Antibiotics: Natural products essential to human health. Med. Res. Rev. 29, 821–842. doi:10.1002/med.20154
Dhillon, B.K., Laird, M.R., Shay, J. a., Winsor, G.L., Lo, R., Nizam, F., Pereira, S.K., Waglechner, N., McArthur, a. G., Langille, M.G.I., Brinkman, F.S.L., 2015. IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. Nucleic Acids Res. 1–5. doi:10.1093/nar/gkv401
Diaz-Torres, M.L., McNab, R., Spratt, D. a., Villedieu, a., Hunt, N., Wilson, M., Mullany, P., 2003. Novel tetracycline resistance determinant from the oral metagenome. Antimicrob. Agents Chemother. 47, 1430–1432. doi:10.1128/AAC.47.4.1430-1432.2003
Dobrindt, U., Hacker, J., 2001. Whole genome plasticity in pathogenic bacteria. Curr. Opin. Microbiol. 4, 550–557. doi:10.1016/S1369-5274(00)00250-2
Drlica, K., Hiasa, H., Kerns, R., Malik, M., Mustaev, A., Zhao, X., 2009. Quinolones: action and resistance updated. Curr. Top. Med. Chem. 9, 981–998. doi:10.2174/156802609789630947
Drury, B., Scott, J., Rosi-Marshall, E.J., Kelly, J.J., 2013. Triclosan exposure increases triclosan resistance and influences taxonomic composition of benthic bacterial communities. Environ. Sci. Technol. 47, 8923–30. doi:10.1021/es401919k
Dutta, C., Pan, A., 2002. Horizontal gene transfer and bacterial diversity. J. Biosci. 27, 27–33. doi:10.1007/BF02703681

Engberg, J., Aarestrup, F.M., Taylor, D.E., Gerner-Smidt, P., Nachamkin, I., 2001. Quinolone and macrolide resistance in Campylobacter jejuni and C. coli: Resistance mechanisms and trends in human isolates. Emerg. Infect. Dis. 7, 24–34. doi:10.3201/eid0701.010104

Fajardo, a., Linares, J.F., Martínez, J.L., 2009. Towards an ecological approach to antibiotics and antibiotic resistance genes. Clin. Microbiol. Infect. 15, 14–16. doi:10.1111/j.1469-0691.2008.02688.x

Fan, F., Yan, K., Wallis, N.G., Reed, S., Moore, T.D., Rittenhouse, S.F., Walter, E., Jr, D., Huang, J., Mcdevitt, D., William, H., Seefeld, M. a, Newlander, K. a, Dalia, R., Head, M.S., Payne, D.J., Dewolf, W.E., Miller, W.H., Jakas, D.R., 2002. Defining and Combating the Mechanisms of Triclosan Resistance in Clinical Isolates of Staphylococcus aureus Defining and Combating the Mechanisms of Triclosan Resistance in Clinical Isolates of Staphylococcus aureus 46, 3343–3347. doi:10.1128/AAC.46.11.3343

Farias, P., Espírito Santo, C., Branco, R., Francisco, R., Santos, S., Hansen, L., Sorensen, S., Morais, P. V., 2015. Natural hotspots for multiple resistances gain: arsenic and antibiotic resistance in heterotrophic aerobic bacteria from marine hydrothermal vent fields. Appl. Environ. Microbiol. AEM.03240–14. doi:10.1128/AEM.03240-14

Forsman, a, 2014. Rethinking phenotypic plasticity and its consequences for individuals, populations and species. Heredity (Edinb). 1–9. doi:10.1038/hdy.2014.92

Fouhy, Y., Scanlon, K., Schouest, K., Spillane, C., Crossman, L., Avison, M., Ryan, R.P., Dow, J.M., 2007. Diffusible signal factor-dependent cell-cell signaling and virulence in the nosocomial pathogen Stenotrophomonas maltophilia. J. Bacteriol. 189, 4964–4968. doi:10.1128/JB.00310-07

Fujita, J., Yamadori, I., Xu, G., Hojo, S., Negayama, K., Miyawaki, H., Yamaji, Y., Takahara, J., 1996. Clinical features of Stenotrophomonas maltophilia pneumonia in immunocompromised patients. Respir. Med. 90, 35–38. doi:10.1016/S0954-6111(96)90242-5

Furuya, E.Y., Lowy, F.D., 2006. Antimicrobial-resistant bacteria in the community setting. Nat. Rev. Microbiol. 4, 36–45. doi:10.1038/nrmicro1325

Gao, S., Seo, J.S., Wang, J., Keum, Y.S., Li, J., Li, Q.X., 2013. Multiple degradation pathways of phenanthrene by Stenotrophomonas maltophilia C6. Int. Biodeterior. Biodegrad. 79, 98–104. doi:10.1016/j.ibiod.2013.01.012

García, P., Monjardín, C., Martín, R., Madera, C., Soberón, N., Garcia, E., Meana, A., Suárez, J.E., 2008. Isolation of new Stenotrophomonas bacteriophages and genomic characterization of temperate phage S1. Appl. Environ. Microbiol. 74, 7552–60. doi:10.1128/AEM.01709-08

Gherardi, G., Creti, R., Pompilio, A., Di Bonaventura, G., 2015. An overview of various typing methods for clinical epidemiology of the emerging pathogen Stenotrophomonas maltophilia. Diagn. Microbiol. Infect. Dis. 81, 219–226. doi:10.1016/j.diagmicrobio.2014.11.005

Gillings, M.R., 2013. Evolutionary consequences of antibiotic use for the resistome, mobilome, and microbial pangenome. Front. Microbiol. 4, 1–10. doi:10.3389/fmicb.2013.00004

Gillings, M.R., Stokes, H.W., 2012. Are humans increasing bacterial evolvability? Trends Ecol. Evol. 27, 346–352. doi:10.1016/j.tree.2012.02.006

Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S.R., Bateman, A., 2005. Rfam: Annotating non-coding RNAs in complete genomes. Nucleic Acids Res. 33, 121–124. doi:10.1093/nar/gki081

Halden, R.U., Heidler, J., Sapkota, A., 2006. Partitioning, persistence, and accumulation in digested sludge of the topical antiseptic triclocarban during wastewater treatment. Environ. Sci. Technol. 40, 3634–9.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., Goodman, R.M., 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem. Biol. 5, R245–9.

Handelsman, J., Tiedje, J., 2007. THE NEW SCIENCE OF METAGENOMICS: Revealing the Secrets of Our Microbial Planet, THE NATION. ed, National Academies Press. The National Academies Press, Washington, DC. doi:NBK54006 [bookaccession]

Härtel, C., Scholz, T., Kuhn, M., Bendiks, M., Göpel, W., Lauten, M., Herting, E., 2012. Innate immune responses to Stenotrophomonas maltophilia in immunocompromised pediatric patients and the effect of taurolidine. J. Microbiol. Immunol. Infect. 115–120. doi:10.1016/j.jmii.2012.04.002

Healy, F.G., Ray, R.M., Aldrich, H.C., Wilkie, a. C., Ingram, L.O., Shanmugam, K.T., 1995. Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic, anaerobic digester maintained on lignocellulose. Appl. Microbiol. Biotechnol. 43, 667–674. doi:10.1007/s002530050468

Hilker, R., Munder, A., Klockgether, J., Losada, P.M., Chouvarine, P., Cramer, N., Davenport, C.F., Dethlefsen, S., Fischer, S., Peng, H., Schönfelder, T., Türk, O., Wiehlmann, L., Wölbeling, F., Gulbins, E., Goesmann, A., Tümmler, B., 2014. Interclonal gradient of virulence in the Pseudomonas aeruginosa pangenome from disease and environment. Environ. Microbiol. doi:10.1111/1462-2920.12606

Howe, R. a., Wilson, M.P., Walsh, T.R., Millar, M.R., 1997. Susceptibility testing of Stenotrophomonas maltophilia to carbapenems. J. Antimicrob. Chemother. 40, 13–17. doi:10.1093/jac/40.1.13

Huedo, P., Conchillo-Solé, O., Yero, D., Martínez-Servat, S., Daura, X., Gibert, I., 2014. Draft Genome Sequence of Stenotrophomonas maltophilia Strain M30, Isolated from a Chronic Pressure Ulcer in an Elderly Patient. Genome Announc. 2, 2009–2010. doi:10.1128/genomeA.00576-14

Huedo, P., Yero, D., Martinez-Servat, S., Estibariz, I., Planell, R., Martinez, P., Ruyra, a., Roher, N., Roca, I., Vila, J., Daura, X., Gibert, I., 2014. Two Different rpf Clusters Distributed among a Population of Stenotrophomonas maltophilia Clinical Strains Display Differential Diffusible Signal Factor Production and Virulence Regulation. J. Bacteriol. 196, 2431–2442. doi:10.1128/JB.01540-14

Hugh, R., Ryschenkow, E., 1961. Pseudomonas maltophilia, an alcaligenes-like species. J. Gen. Microbiol. 26, 123–132. doi:10.1099/00221287-26-1-123

Jackson, R.W., Vinatzer, B., Arnold, D.L., Dorus, S.S., Murillo, J., 2011. The influence of the accessory genome on bacterial pathogen evolution. Mob. Genet. Elements 1, 55–65. doi:10.4161/mge.1.1.16432

Jakobsson, H.E., Jernberg, C., Andersson, A.F., Sjölund-Karlsson, M., Jansson, J.K., Engstrand, L., 2010. Short-term antibiotic treatment has differing long- term impacts on the human throat and gut microbiome. PLoS One 5. doi:10.1371/journal.pone.0009836

Jernberg, C., Löfmark, S., Edlund, C., Jansson, J.K., 2010. Long-term impacts of antibiotic exposure on the human intestinal microbiota. Microbiology 156, 3216–3223. doi:10.1099/mic.0.040618-0

Kaas, R.S., Friis, C., Ussery, D.W., Aarestrup, F.M., 2012. Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse Escherichia coli genomes. BMC Genomics 13, 577. doi:10.1186/1471-2164-13-577

Kalyuzhnaya, M.G., De Marco, P., Bowerman, S., Pacheco, C.C., Lara, J.C., Lidstrom, M.E., Chistoserdova, L., 2006. Methyloversatilis universalis gen. nov., sp. nov., a novel taxon within the Betaproteobacteria represented by three methylotrophic isolates. Int. J. Syst. Evol. Microbiol. 56, 2517–2522. doi:10.1099/ijs.0.64422-0

Knapp, J.S., Fox, K.K., Trees, D.L., Whittington, W.L., 1997. Fluoroquinolone resistance in Neisseria gonorrhoeae. Emerg. Infect. Dis. 3, 33–39. doi:10.3201/eid0301.970104

Kohanski, M. a, Dwyer, D.J., Collins, J.J., 2010. How antibiotics kill bacteria: from targets to networks. Nat. Rev. Microbiol. 8, 423–435. doi:10.1038/nrmicro2333

Kolpin, D.W., Furlong, E.T., Meyer, M.T., Thurman, E.M., Zaugg, S.D., Barber, L.B., Buxton, H.T., 2002. Pharmaceuticals , hormones , and other organic wastewater Contaminants in U . S . Streams , 1999 - 2000 : A National Reconnaissance. Environ. Sci. Technol. 36, 1202–11. doi:10.1021/es011055j

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., Schloss, P.D., 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. Appl. Environ. Microbiol. 79, 5112–5120. doi:10.1128/AEM.01043-13

Kristensen, D.M., Kannan, L., Coleman, M.K., Wolf, Y.I., Sorokin, A., Koonin, E. V., Mushegian, A., 2010. A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. Bioinformatics 26, 1481–1487. doi:10.1093/bioinformatics/btq229

Kümmerer, K., 2004. Resistance in the environment. J. Antimicrob. Chemother. 54, 311–20. doi:10.1093/jac/dkh325

Kung, V.L., Ozer, E. a, Hauser, A.R., 2010. The accessory genome of Pseudomonas aeruginosa. Microbiol. Mol. Biol. Rev. 74, 621–41. doi:10.1128/MMBR.00027-10
Lascols, C., Podglajen, I., Verdet, C., Gautier, V., Gutmann, L., Soussy, C.J., Collatz, E., Cambau, E., 2008. A plasmid-borne Shewanella algae gene, qnrA3, and its possible transfer in vivo between Kluyvera ascorbata and Klebsiella pneumoniae. J. Bacteriol. 190, 5217–5223. doi:10.1128/JB.00243-08
Lawrence, J.G., Ochman, H., Ragan, M. a., 2002. Reconciling the many faces of lateral gene transfer. Trends Microbiol. 10, 1–4. doi:10.1016/S0966-842X(01)02282-X
Ledder, R.G., Gilbert, P., Willis, C., McBain, a. J., 2006. Effects of chronic triclosan exposure upon the antimicrobial susceptibility of 40 ex-situ environmental and human isolates. J. Appl. Microbiol. 100, 1132–1140. doi:10.1111/j.1365-2672.2006.02811.x
Leplae, R., Hebrant, A., Wodak, S.J., Toussaint, A., 2004. ACLAME: a classification of mobile genetic elements. Nucleic Acids Res. 32, D45–D49. doi:10.1093/nar/gkh084
Leplae, R., Lima-Mendez, G., Toussaint, A., 2010. ACLAME: a Classification of mobile genetic elements, update 2010. Nucleic Acids Res. 38, D57–61. doi:10.1093/nar/gkp938
Levy, S.B., 2002. Factors impacting on the problem of antibiotic resistance. J. Antimicrob. Chemother. 49, 25–30. doi:10.1093/jac/49.1.25
Lira, F., Hernández, A., Belda, E., Sánchez, M.B., Moya, A., Silva, F.J., Martínez, J.L., 2012. Whole-genome sequence of Stenotrophomonas maltophilia D457, a clinical isolate and a model strain. J. … 194, 3563–4. doi:10.1128/JB.00602-12
Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964. doi:10.1093/nar/25.5.955
Lucas, S., Copeland, A., Lapidus, A., Glavina del Rio, T., Dalin, E., Tice, H., Pitluck, S., Chain, P., Malfatti, S., Shin, M., Vergez, L., Lang, D., Schmutz, J., Larimer, F., Land, M., Hauser, L., Kyrpides, N., Mikhailova, N., Taghavi, S., Monchy, S., Newman, L., Vangronsveld, J., van der Lelie, D., Richardson, P., 2008. Complete sequence of Stenotrophomonas maltophilia R551-3. Unpublished JGI Projec.
Lucas, S., Han, J., Lapidus, A., Cheng, J.F., Goodwin, L., Pitluck, S., Peters, L., Ovchinnikova, G., Teshima, H., Detter, J.C., Han, C., Tapia, R., Land, M., Hauser, L., Kyrpides, N., Ivanova, N., Pagani, I., Medau, R., Furlan, B., Mui Tsai, S., Rodriques, J., Tiedje, J., Woyke, T., 2011. Complete sequence of Burkholderia sp. JV3.
Lupo, A., Coyne, S., Berendonk, T.U., 2012. Origin and evolution of antibiotic resistance: The common mechanisms of emergence and spread in water bodies. Front. Microbiol. 3, 1–13. doi:10.3389/fmicb.2012.00018
Lutz, R., Bujard, H., 1997. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Res. 25, 1203–10.
Martínez, J.L., 2011. Bottlenecks in the transferability of antibiotic resistance from natural ecosystems to human bacterial pathogens. Front. Microbiol. 2, 265. doi:10.3389/fmicb.2011.00265
Martínez, J.L., Baquero, F., Andersson, D.I., 2011. Beyond serial passages: new methods for predicting the emergence of resistance to novel antibiotics. Curr. Opin. Pharmacol. 11, 439–45. doi:10.1016/j.coph.2011.07.005
Martínez, J.L., Baquero, F., Andersson, D.I., 2007. Predicting antibiotic resistance. Nat. Rev. Microbiol. 5, 958–965. doi:10.1038/nrmicro1796
Martínez-Martínez, L., Pascual, A., Jacoby, G. a., 1998. Quinolone resistance from a transferable plasmid. Lancet 351, 797–799. doi:10.1016/S0140-6736(97)07322-4
Martiny, A.C., Martiny, J.B.H., Weihe, C., Field, A., Ellis, J.C., 2011. Functional metagenomics reveals previously unrecognized diversity of antibiotic resistance genes in gulls. Front. Microbiol. 2, 1–11. doi:10.3389/fmicb.2011.00238
McMurry, L.M., Oethinger, M., Levy, S.B., 1998. Triclosan targets lipid synthesis. Nature 394, 531–532. doi:10.1038/28970
Medini, D., Donati, C., Tettelin, H., Masignani, V., Rappuoli, R., 2005. The microbial pan-genome. Curr. Opin. Genet. Dev. 15, 589–594. doi:10.1016/j.gde.2005.09.006
Mira, A., Martín-Cuadrado, A.B., D'Auria, G., Rodríguez-Valera, F., 2010. The bacterial pan-genome: A new paradigm in microbiology. Int. Microbiol. 13, 45–57. doi:10.2436/20.1501.01.110

Mohamed, M.S., 2009. Degradation of methomyl by the novel bacterial strain Stenotrophomonas maltophilia M1. Electron. J. Biotechnol. 12. doi:10.2225/vol12-issue4-fulltext-11
Mølbak, K., Gerner-Smidt, P., Wegener, H.C., 2002. Increasing quinolone resistance in Salmonella enterica serotype Enteritidis. Emerg. Infect. Dis. 8, 514–515. doi:10.3201/eid0805.010288
Mole, B., 2013. MRSA: Farming up trouble. Nature 499, 398–400. doi:10.1038/499398a
Moore, A.M., Munck, C., Sommer, M.O. a, Dantas, G., 2011. Functional metagenomic investigations of the human intestinal microbiota. Front. Microbiol. 2, 1–8. doi:10.3389/fmicb.2011.00188
Mullany, P., 2014. Functional metagenomics for the investigation of antibiotic resistance. Virulence 5, 443–7. doi:10.4161/viru.28196
Novais, Â., Comas, I., Baquero, F., Cantón, R., Coque, T.M., Moya, A., González-Candelas, F.G., Galán, J.C., 2010. Evolutionary trajectories of beta-lactamase CTX-M-1 cluster enzymes: Predicting antibiotic resistance. PLoS Pathog. 6, 1–16. doi:10.1371/journal.ppat.1000735
Ogram, A., Sayler, G.S., Barkay, T., 1987. The extraction and purification of microbial DNA from sediments. J. Microbiol. Methods 7, 57–66. doi:10.1016/0167-7012(87)90025-X
Ozer, E. a, Allen, J.P., Hauser, A.R., 2014. Characterization of the core and accessory genomes of Pseudomonas aeruginosa using bioinformatic tools Spine and AGEnt. BMC Genomics 15, 1–17. doi:10.1186/1471-2164-15-737
Pace, N., Stahl, D., Lane, D., Olsen, G., 1986. The Analysis of Natural Microbial Populations by Ribosomal RNA Sequences, in: Marshall, K.C. (Ed.), Advances in Microbial Ecology SE - 1, Advances in Microbial Ecology. Springer US, pp. 1–55. doi:10.1007/978-1-4757-0611-6_1
Paez, J.I.G., Tengan, F.M., Barone, a. a., Levin, a. S., Costa, S.F., 2008. Factors associated with mortality in patients with bloodstream infection and pneumonia due to Stenotrophomonas maltophilia. Eur. J. Clin. Microbiol. Infect. Dis. 27, 901–906. doi:10.1007/s10096-008-0518-2
Pages, D., Rose, J., Conrod, S., Cuine, S., 2008. Heavy metal tolerance in Stenotrophomonas maltophilia. PLoS One 3, e1539. doi:10.1371/journal.pone.0001539
Palleroni, N.J., Bradbury, J.F., 1993. Stenotrophomonas, a new bacterial genus for Xanthomonas maltophilia (Hugh 1980) Swings et al. 1983. Int. J. Syst. Bacteriol. 43, 606–609. doi:10.1099/00207713-43-3-606
Palumbi, S.R., 2001. Evolution - Humans as the world´s greatest evolutionary force. Science (80-.). 293, 1786–1790.
Pankuch, G. a., Jacobs, M.R., Rittenhouse, S.F., Appelbaum, P.C., 1994. Susceptibilities of 123 strains of Xanthomonas maltophilia to eight b-lactams (including b-lactam-b-lactamase inhibitor combinations) and ciprofloxacin tested by five methods. Antimicrob. Agents Chemother. 38, 2317–2322. doi:10.1128/AAC.38.10.2317
Patel, R.K., Jain, M., 2012. NGS QC toolkit: A toolkit for quality control of next generation sequencing data. PLoS One 7. doi:10.1371/journal.pone.0030619
Penders, J., Stobberingh, E.E., Savelkoul, P.H.M., Wolffs, P.F.G., 2013. The human microbiome as a reservoir of antimicrobial resistance. Front. Microbiol. 4, 87. doi:10.3389/fmicb.2013.00087
Petersen, A., Andersen, J.S., Somsiri, T., Dalsgaard, A., Kaewmak, T., 2002. Impact of Integrated Fish Farming on Antimicrobial Resistance in a Pond Environment Impact of Integrated Fish Farming on Antimicrobial Resistance in a Pond Environment. doi:10.1128/AEM.68.12.6036
Pompilio, A., Pomponio, S., Crocetta, V., Gherardi, G., Verginelli, F., Fiscarelli, E., Dicuonzo, G., Savini, V., D'Antonio, D., Di Bonaventura, G., 2011. Phenotypic and genotypic characterization of Stenotrophomonas maltophilia isolates from patients with cystic fibrosis: genome diversity, biofilm formation, and virulence. BMC Microbiol. 11, 159. doi:10.1186/1471-2180-11-159
Quinn, J.P., 1998. Clinical problems posed by multiresistant nonfermenting gram-negative pathogens. Clin. Infect. Dis. 27 Suppl 1, S117–S124.
Rasko, D. a., Rosovitz, M.J., Myers, G.S. a, Mongodin, E.F., Fricke, W.F., Gajer, P., Crabtree, J., Sebaihia, M., Thomson, N.R., Chaudhuri, R., Henderson, I.R., Sperandio, V., Ravel, J., 2008. The pangenome structure of Escherichia coli: Comparative genomic analysis of E. coli commensal and pathogenic isolates. J. Bacteriol. 190, 6881–6893. doi:10.1128/JB.00619-08
Reanney, D., 1976. Extrachromosomal elements as possible agents of adaptation and development. Bacteriol. Rev. 40, 552–590.
Redgrave, L.S., Sutton, S.B., Webber, M. a., Piddock, L.J. V, 2014. Fluoroquinolone resistance: Mechanisms, impact on bacteria, and role in evolutionary success. Trends Microbiol. 22, 438–445. doi:10.1016/j.tim.2014.04.007

Reen, F.J., Mooij, M.J., Holcombe, L.J., Mcsweeney, C.M., Mcglacken, G.P., Morrissey, J.P., O'Gara, F., 2011. The Pseudomonas quinolone signal (PQS), and its precursor HHQ, modulate interspecies and interkingdom behaviour. FEMS Microbiol. Ecol. 77, 413–428. doi:10.1111/j.1574-6941.2011.01121.x

Rice, L.B., 2002. Association of different mobile elements to generate novel integrative elements. Cell. Mol. Life Sci. 59, 2023–2032. doi:10.1007/s000180200002

Rizzo, L., Manaia, C.M., Merlin, C., Schwartz, T., Dagot, C., Ploy, M.C., Michael, I., Fatta-Kassinos, D., 2013. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: a review. Sci. Total Environ. 447, 345–60. doi:10.1016/j.scitotenv.2013.01.032

Romanenko, L. a., Uchino, M., Tanaka, N., Frolova, G.M., Slinkina, N.N., Mikhailov, V. V., 2008. Occurrence and antagonistic potential of Stenotrophomonas strains isolated from deep-sea invertebrates. Arch. Microbiol. 189, 337–344. doi:10.1007/s00203-007-0324-8

Rowe-Magnus, D. a, Mazel, D., 2002. The role of integrons in antibiotic resistance gene capture. Int. J. Med. Microbiol. 292, 115–125. doi:10.1078/1438-4221-00197

Russell, a D., 2000. Do biocides select for antibiotic resistance? J. Pharm. Pharmacol. 52, 227–33.

Ryan, R.P., Monchy, S., Cardinale, M., Taghavi, S., Crossman, L., Avison, M.B., Berg, G., van der Lelie, D., Dow, J.M., 2009. The versatility and adaptation of bacteria from the genus Stenotrophomonas. Nat. Rev. Microbiol. 7, 514–525. doi:10.1038/nrmicro2163

Sánchez, M.B., Martínez, J.L., 2012. Differential epigenetic compatibility of qnr antibiotic resistance determinants with the chromosome of Escherichia coli. PLoS One 7. doi:10.1371/journal.pone.0035149

Sanders, C.C., 1989. ??-Lactamase stability and in vitro activity of oral cephalosporins against strains possessing well-characterized mechanisms of resistance. Antimicrob. Agents Chemother. 33, 1313–1317. doi:10.1128/AAC.33.8.1313

Schmieder, R., Edwards, R., 2012. Insights into antibiotic resistance through metagenomic approaches. Future Microbiol. 7, 73–89. doi:10.2217/fmb.11.135

Schweizer, H.P.H.P., 2001. Triclosan: a widely used biocide and its link to antibiotics. FEMS Microbiol. Lett. 202, 1–7. doi:10.1016/S0378-1097(01)00273-7

Seemann, T., 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. doi:10.1093/bioinformatics/btu153

Shapiro, J. a., 2005. A 21st century view of evolution: Genome system architecture, repetitive DNA, and natural genetic engineering. Gene 345, 91–100. doi:10.1016/j.gene.2004.11.020

Siefert, J.L., 2009a. Horizontal Gene Transfer. Methods in Molecular Biology 532, 13–27. doi:10.1007/978-1-60327-853-9

Siefert, J.L., 2009b. Defining the mobilome. Methods Mol. Biol., Methods in Molecular Biology 532, 13–27. doi:10.1007/978-1-60327-853-9_2

Silva, F.J., Belda, E., Talens, S.E., 2006. Differential annotation of tRNA genes with anticodon CAT in bacterial genomes. Nucleic Acids Res. 34, 6015–6022. doi:10.1093/nar/gkl739

Smith, J.M., 1970. Natural selection and the concept of a protein space. Nature 225, 563–564. doi:10.1038/225563a0

Smith, M.A., Bidochka, M.J., 1998. Bacterial fitness and plasmid loss: the importance of culture conditions and plasmid size. Can. J. Microbiol. 44, 351–355. doi:10.1139/w98-020

Sommer, M.O. a, Dantas, G., 2011. Antibiotics and the resistant microbiome. Curr. Opin. Microbiol. 14, 556–63. doi:10.1016/j.mib.2011.07.005

Sommer, M.O. a, Dantas, G., Church, G.M., 2009. Functional characterization of the antibiotic resistance reservoir in the human microflora. Science 325, 1128–31. doi:10.1126/science.1176950

Staden, R., Beal, K.F., Bonfield, J.K., 2000. The Staden package, 1998. Methods Mol. Biol. 132, 115–130. doi:10.1385/1-59259-192-2:115

Stokes, H.W., Gillings, M.R., 2011. Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens. FEMS Microbiol. Rev. 35, 790–819. doi:10.1111/j.1574-6976.2011.00273.x

Strahilevitz, J., Jacoby, G., 2009. Plasmid-mediated quinolone resistance: a multifaceted threat. Clin. Microbiol. … 22, 664–689. doi:10.1128/CMR.00016-09

Suckstorff, I., Berg, G., 2003. Evidence for dose-dependent effects on plant growth by Stenotrophomonas strains from different origins. J. Appl. Microbiol. 95, 656–663. doi:10.1046/j.1365-2672.2003.02021.x

Swings, J., De Vos, P., Van den Mooter, M., De Ley, J., 1983. Transfer of Pseudomonas maltophilia Hugh 1981 to the Genus Xanthomonas as Xanthomonas maltophilia (Hugh 1981) comb. nov. Int. J. Syst. Bacteriol. doi:10.1099/00207713-33-2-409

Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S. V, Crabtree, J., Jones, A.L., Durkin, a S., Deboy, R.T., Davidsen, T.M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J.D., Hauser, C.R., Sundaram, J.P., Nelson, W.C., Madupu, R., Brinkac, L.M., Dodson, R.J., Rosovitz, M.J., Sullivan, S. a, Daugherty, S.C., Haft, D.H., Selengut, J., Gwinn, M.L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'Connor, K.J.B., Smith, S., Utterback, T.R., White, O., Rubens, C.E., Grandi, G., Madoff, L.C., Kasper, D.L., Telford, J.L., Wessels, M.R., Rappuoli, R., Fraser, C.M., 2005. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome". Proc. Natl. Acad. Sci. U. S. A. 102, 13950–13955. doi:10.1073/pnas.0506758102

Tettelin, H., Riley, D., Cattuto, C., Medini, D., 2008. Comparative genomics: the bacterial pan-genome. Curr. Opin. Microbiol. 11, 472–477. doi:10.1016/j.mib.2008.09.006

Thomas, V.L., McReynolds, A.C., Shoichet, B.K., 2010. Structural Bases for Stability-Function Tradeoffs in Antibiotic Resistance. J. Mol. Biol. 396, 47–59. doi:10.1016/j.jmb.2009.11.005

Torres-Cortés, G., Millán, V., Ramírez-Saad, H.C., Nisa-Martínez, R., Toro, N., Martínez-Abarca, F., 2011. Characterization of novel antibiotic resistance genes identified by functional metagenomics on soil samples. Environ. Microbiol. 13, 1101–14. doi:10.1111/j.1462-2920.2010.02422.x

Turnbaugh, P.J., Ley, R.E., Hamady, M., Fraser-Liggett, C.M., Knight, R., Gordon, J.I., 2007. The Human Microbiome Project. Nature 449, 804–810. doi:10.1038/nature06244

Van Tonder, A.J., Mistry, S., Bray, J.E., Hill, D.M.C., Cody, A.J., Farmer, C.L., Klugman, K.P., von Gottberg, A., Bentley, S.D., Parkhill, J., Jolley, K.A., Maiden, M.C.J., Brueggemann, A.B., 2014. Defining the Estimated Core Genome of Bacterial Populations Using a Bayesian Decision Model. PLoS Comput. Biol. 10, e1003788. doi:10.1371/journal.pcbi.1003788

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J. a, Wu, D., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y.-H., Smith, H.O., 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science 304, 66–74. doi:10.1126/science.1093857

Vernikos, G., Medini, D., Riley, D.R., Tettelin, H., 2015. Ten years of pan-genome analyses. Curr. Opin. Microbiol. 23, 148–154. doi:10.1016/j.mib.2014.11.016

Waldor, M.K., 2010. Mobilizable genomic islands: Going mobile with oriT mimicry. Mol. Microbiol. 78, 537–540. doi:10.1111/j.1365-2958.2010.07365.x

Walsh, F., 2013. Investigating antibiotic resistance in non-clinical environments. Front. Microbiol. 4, 1–5. doi:10.3389/fmicb.2013.00019

Wang, X., Chen, J., Quinn, P., 2012. Reprogramming Microbial Metabolic Pathways, Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki. doi:10.1007/978-94-007-5055-5

Weinreich, D.M., Delaney, N.F., Depristo, M. a, Hartl, D.L., 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science 312, 111–114. doi:10.1126/science.1123539

Wilmes, P., Simmons, S.L., Denef, V.J., Banfield, J.F., 2009. The dynamic genetic repertoire of microbial communities. FEMS Microbiol. Rev. 33, 109–132. doi:10.1111/j.1574-6976.2008.00144.x

Woodford, N., Ellington, M.J., 2007. The emergence of antibiotic resistance by mutation. Clin. Microbiol. Infect. 13, 5–18. doi:10.1111/j.1469-0691.2006.01492.x

World Health Organization, 2014. ANTIMICROBIAL RESISTANCE Global Report on Surveillance 2014. World Heal. Organ.

World Health Organization, 2002. Use of Antimicrobials Outside Human Medicine and Resultant Antimicrobial Resistance in Humans,Highlights of Fact Sheet No. 268 [WWW Document]. Fact Sheet No. 268. URL http://whqlibdoc.who.int/fact_sheet/2002/FS_268.pdf

Wren, B.W., 2000. Microbial genome analysis: insights into virulence, host adaptation and evolution. Nat. Rev. Genet. 1, 30–39. doi:10.1038/35049551

Wright, G., 2014. Perspective: Synthetic biology revives antibiotics. Nature 509, S13–S13. doi:10.1038/509S13a

Wright, G.D., 2010. The antibiotic resistome. Expert Opin. Drug Discov. 5, 779–88. doi:10.1517/17460441.2010.497535

Xu, X., Liu, L., Jia, Z., Shu, Y., 2015. Determination of enrofloxacin and ciprofloxacin in foods of animal origin by capillary electrophoresis with field amplified sample stacking–sweeping technique. Food Chem. 176, 219–225. doi:10.1016/j.foodchem.2014.12.054

Yang, H., Chen, H., Yang, Q., Chen, M., Wang, H., 2008. High prevalence of plasmid-mediated quinolone resistance genes qnr and aac(6')-Ib-cr in clinical isolates of Enterobacteriaceae from nine teaching hospitals in China. Antimicrob. Agents Chemother. 52, 4268–4273. doi:10.1128/AAC.00830-08

Yazdankhah, S.P., Scheie, A. a, Høiby, E.A., Lunestad, B.-T., Heir, E., Fotland, T.Ø., Naterstad, K., Kruse, H., 2006. Triclosan and antimicrobial resistance in bacteria: an overview. Microb. Drug Resist. 12, 83–90. doi:10.1089/mdr.2006.12.83

Zhang, L., Li, X.Z., Poole, K., 2001. SmeDEF multidrug efflux pump contributes to intrinsic multidrug resistance in Stenotrophomonas maltophilia. Antimicrob. Agents Chemother. 45, 3497–503. doi:10.1128/AAC.45.12.3497-3503.2001

Zhang, R., Sun, Q., Hu, Y.J., Yu, H., Li, Y., Shen, Q., Li, G.X., Cao, J.M., Yang, W., Wang, Q., Zhou, H.W., Hu, Y.Y., Chen, G.X., 2012. Detection of the Smqnr quinolone protection gene and its prevalence in clinical isolates of Stenotrophomonas maltophilia in China. J. Med. Microbiol. 61, 535–539. doi:10.1099/jmm.0.037309-0

Zhang, Z., Chisti, Y., 1996. Plasmid stability in recombinant sa ccharomyces cerevisiae. Science (80-.). 14, 401–435.

Zhao, Y., Tang, H., Ye, Y., 2012. RAPSearch2: A fast and memory-efficient protein similarity search tool for next-generation sequencing data. Bioinformatics 28, 125–126. doi:10.1093/bioinformatics/btr595

Zhou, W., Wang, Y., Lin, J., 2012. Functional cloning and characterization of antibiotic resistance genes from the chicken gut microbiome. Appl. Environ. Microbiol. 78, 3028–32. doi:10.1128/AEM.06920-11

Zhu, B., Liu, H., Tian, W.-X.X., Fan, X.-Y.Y., Li, B., Zhou, X.-P.P., Jin, G.-L.L., Xie, G.-L.L., 2012. Genome sequence of Stenotrophomonas maltophilia RR-10, isolated as an endophyte from rice root. J. Bacteriol. 194, 1280–1281. doi:10.1128/JB.06702-11

Appendix

9. Appendix

Appendix 1: Summary of the pre- and post-processing data of twenty isolates of *Stenotrophomonas maltophilia*.

Code	Isolates	G+C%	Total reads	Reads filtered (QC)	Reads assembled	N of contigs	Total consensus	Largest contigs	N50 contig size	N90 contigs size	N95 contigs size	CDS	RNA
FL1	E729	66	3471954	2428220(69.94%)	2361177(97.24%)	175	5077081	280989	57021	16939	11197	4690	76
FL2	E759	66	4425476	3007154(67.95%)	2929889(97.43%)	108	4597078	290826	82967	19346	13258	4164	71
FL3	E999	67	4087678	2526112(61.8%)	2457344(97.28%)	156	4481504	196468	60419	15289	8895	4007	58
FL4	G51	66	4754284	3254918(68.46%)	3163464(97.19%)	178	4933493	280033	56532	14446	8350	4499	63
FL5	E301	67	4797882	3330888 (69.42%)	3244331(97.4%)	127	4480288	253397	70621	20388	12093	4050	72
FL6	D388	66	6492918	4316864(66.49%)	4209715(97.52%)	159	4766556	252083	65976	16477	9611	4316	73
FL7	E861	66	4951424	3372866(68.12%)	3279000(97.22%)	207	4762067	132329	45010	10222	6674	4319	73
FL8	C357	66	4984782	3300798(66.22%)	3225416(97.71%)	146	4874541	165749	59737	17046	11221	4415	68
FL9	E539	66	4228078	2897382(68.53%)	2809525(96.97%)	204	4661407	115973	46377	11228	7021	4196	64
FL10	E824	66	4524164	3155564(69.75%)	3061904(97.03%)	359	5207561	128802	33215	7703	4016	4728	60
FL11	N-S26	66	3463586	2097382(60.55%)	2048110(97.65%)	129	4749051	156334	91786	18738	12621	4242	74
FL12	EP13	66	4794388	3352074(69.92%)	3267743(97.48%)	183	4848230	167604	61157	15624	8349	4412	71
FL13	EA22	66	4656134	3197076(68.66%)	3115246(97.44%)	144	4832045	211509	67660	23015	10347	4358	77
FL14	EA1	66	3353214	2366232(70.57%)	2300858(97.24%)	187	4835161	178120	45151	16788	8565	4343	74
FL15	PS5	66	4510776	2900984(64.31%)	2816725(97.1%)	145	4661735	229479	52609	17288	10592	4183	65
FL16	EA23	66	4014038	2859010(71.23%)	2784940(97.41%)	198	4849545	153482	52500	15005	10046	4424	73
FL17	EP20	66	3644638	2515324(69.01%)	2452933(97.52%)	158	4694885	178099	47334	16275	11051	4205	70
FL18	EP5	66	3806172	2698572(70.90%)	2628194(97.39%)	162	4660244	258776	55493	14503	9154	4180	58
FL19	EA21	66	4403242	3021878(68.63%)	2932974(97.06%)	156	4804616	198620	60961	15931	10069	4346	72
FL20	EA63	66	3607036	2434350(67.49%)	2369787(97.35%)	124	4946929	209470	84640	28331	18294	4471	77

Appendix 2: Resistance genes and putative resistance genes detected at all strains analyzed at this study. All genes were collected from *S. maltophilia* D457 genome as reference. In cases where the reference genes were obtained from *S. maltophilia* K279a (*) or *S. maltophilia* R551-3 (**), they are indicate.

Reference D457	Genes	K279	D457	JV3	R551-3	FL1	FL2	FL3	FL4	FL5	FL6	FL7	FL8	FL9	FL10	FL11	FL12	FL13	FL14	FL15	FL16	FL17	FL18	FL19	FL20
GI:504458362	Beta lactamase resistance - *ampC*	■	■	■	■	■	□	■	■	□	■	■	■	■	■	■	■	■	□	■	■	■	■	■	■
GI:754362349	Beta lactamase resistance - *bla1*	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:504461071	Beta lactamase resistance - *bla2*	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:758874050*	Know and putative aminoglycoside resistance gene: Aminoglycoside 6N-acetiltransferase (aac(6')-iz)	■	□	□	□	□	□	□	□	□	□	□	□	□	□	□	■	□	□	□	■	□	□	□	□
GI:504458884	Know and putative aminoglycoside resistance gene: Dimethiladenosine transferase (ksgA)	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:765003578*	Know and putative aminoglycoside resistance gene: Putative aminoglycoside 2'-N-acetiiltransferase	■	□	□	□	□	■	□	■	□	□	□	□	□	□	□	□	■	■	□	□	□	□	■	■
GI:504459811	Know and putative aminoglycoside resistance gene: Putative aminoglycoside 3'-phosphotransferase	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■
GI:504458430	Know and putative aminoglycoside resistance gene: Putative aminoglycoside phosphotransferase	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:504459147	Know and putative aminoglycoside resistance gene: Putative aminoglycoside phosphotransferase	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:504459814	Know and putative aminoglycoside resistance gene: Putative spectinomycin phosphotransferase	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■	■	■
GI:504459979	Know and putative aminoglycoside resistance gene: Spectinomycin 3'-phosphotransferase	■	■	■	■	■	□	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■	□	■
GI:504459101	Fluoroquinolone resistance protein qnrB (smqnr)	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:754361851	Phosphomannomutase/phosphoglucomutase (spgM)	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■	■	■	■	■	■	■
GI:5015041180**	Putative B-lactam resistance genes: Beta-lactamase	□	□	□	■	□	□	□	□	■	□	□	■	□	■	■	□	□	□	□	□	□	□	□	□
GI:754363045	Putative B-lactam resistance genes: Beta-lactamase class C	□	■	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
GI:504458531	Putative B-lactam resistance genes: Metalo-beta-lactamase family protein	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■	■	■	■	■	■
GI:504460542	Putative B-lactam resistance genes: Putative beta-lactamase	■	■	□	■	■	■	□	□	□	□	□	■	□	■	■	■	□	■	■	■	□	□	■	□
GI:504460526	Putative B-lactam resistance genes: Putative beta-lactamase	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■	■	■	■	■	■	■
GI:754362052	Putative B-lactam resistance genes: Putative beta-lactamase	■	■	■	■	■	■	■	■	■	■	□	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:504458662	Putative B-lactam resistance genes: Putative beta-lactamase	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■	■	■	■	■
GI:504460856	Putative B-lactam resistance genes: Putative beta-lactamase	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:504460542	Putative B-lactam resistance genes: Putative beta-lactamase	■	■	□	■	■	■	□	□	□	□	□	■	□	■	■	■	□	■	■	■	□	□	■	□
GI:501457908*	Putative B-lactam resistance genes: Putative beta-lactamase	■	□	□	□	□	□	□	■	□	■	■	□	■	□	□	■	■	□	□	■	□	□	■	■
GI:501456828*	Putative B-lactam resistance genes: Putative beta-lactamase	■	□	□	■	□	■	■	■	□	□	□	■	■	□	□	■	■	□	■	■	■	■	■	■
GI:754362890	Putative B-lactam resistance genes: Putative beta-lactamase (PBP4)	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■	■	■	■	■	■	■	■	■	■
GI:501456881*	Putative B-lactam resistance genes: Putative beta-lactamase AmpC protein	■	□	□	□	■	□	□	□	■	■	■	■	■	□	□	■	□	■	■	■	■	■	□	□
GI:504461490	Putative B-lactam resistance genes: Putative beta-lactamase protein	■	■	■	■	■	■	■	■	■	■	■	□	■	□	□	■	■	■	■	■	□	■	■	■
GI:504460976	Putative B-lactam resistance genes: Putative beta-lactamase protein	■	■	■	■	■	■	■	■	□	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:504458711	Putative B-lactam resistance genes: Putative metallo-beta-lactamase family protein	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:504458712	Putative B-lactam resistance genes: Putative metallo-beta-lactamase family protein	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:504461297	Putative B-lactam resistance genes: Putative metallo-beta-lactamase superfamily protein	■	■	■	■	■	■	■	■	□	■	■	■	■	■	■	■	■	■	□	■	■	■	■	■
GI:518168129	Putative B-lactam resistance genes: Putative metallo-beta-lactamase superfamily protein	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:504461139	Putative B-lactam resistance genes: Putative metallo-beta-lactamase superfamily protein	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GI:754362316	Putative B-lactam resistance genes: Putative penicillin-binding protein/beta-lactamase	■	■	□	□	■	■	■	■	□	■	■	□	■	□	□	■	■	□	■	■	■	■	■	■

Appendix 3: Genomic islands (GEI) present at the genome of *S. maltophilia* D457 obtained with the software islandViewer.

GEI	Start	End	Size	Locus ID	Gene name	Product
GEI-1	54822	63101	8279	SMD_0047		hypothetical protein
				SMD_0048	C4.3	c4 antisense RNA (Rfam family RF01695)
				SMD_0049		wall associated protein
				SMD_0050		wall associated protein
GEI-2	162624	201509	38885	SMD_0137		hypothetical protein
				SMD_0138		membrane-bound lytic murein transglycosylase
				SMD_0139	ISD1-TranspB	ISStmaD1 Transposase B
				SMD_0140	ISD1-TranspA	ISStmaD1 Transposase A
				SMD_0142	smmP2	Cobalt/zinc/cadmium efflux RND transporter,membrane fusion protein, CzcB family
				SMD_0143	smmQ2	Cobalt-zinc-cadmium resistance protein CzcA Cation efflux system protein CusA
				SMD_0144	nreB	NreB protein
				SMD_0145		hypothetical protein
				SMD_0146		hypothetical protein
				SMD_0147	smmJ	heavy metal RND efflux outer membrane protein,CzcC family
				SMD_0148	smmK	Co/Zn/Cd efflux system membrane fusion protein
				SMD_0149	cusA	Cobalt-zinc-cadmium resistance protein CzcA Cation efflux system protein CusA
				SMD_0150		hypothetical protein
				SMD_0152		hypothetical protein
				SMD_0153		hypothetical protein
				SMD_0154	araC reg	AraC family transcriptional regulator
				SMD_0155		glyoxalase
				SMD_0156		Aspartate aminotransferase
				SMD_0157		hypothetical protein
				SMD_0158	pnuC	PnuC-like transporter linked to homoserine kinase and OMR
				SMD_0159		ferrichrome-iron receptor
				SMD_0160		homoserine kinase type II,PnuC-associated, THI-regulated branch
				SMD_0138		membrane-bound lytic murein transglycosylase
				SMD_0139	ISD1-TranspB	ISStmaD1 Transposase B
				SMD_0140	ISD1-TranspA	ISStmaD1 Transposase A
				SMD_0142	smmP2	Cobalt/zinc/cadmium efflux RND transporter,membrane fusion protein, CzcB family
				SMD_0143	smmQ2	Cobalt-zinc-cadmium resistance protein CzcA Cation efflux system protein CusA
GEI-3	293704	299236	5532	SMD_0242		putative minor tail protein
				SMD_0243		hypothetical protein
				SMD_0244		deduced tail fiber protein
				SMD_0245		deduced tail fiber protein
				SMD_0246		hypothetical protein
				SMD_0247		Mu-like phage tail fiber protein
GEI-4	317692	324915	7223	SMD_0265	birA	biotin--protein ligase
				SMD_0266	coaX	Pantothenate kinase type III, CoaX-like
				SMD_0267		hypothetical protein
				SMD_0269	ISD1-TranspB	ISStmaD1 Transposase B
				SMD_0270	ISD1-TranspA	ISStmaD1 Transposase A
				SMD_0271		hypothetical protein
				SMD_0272		transcriptional regulator
				SMD_0273		hypothetical protein
GEI-5	976021	1019386	43365	SMD_0861		hypothetical protein
				SMD_0867		hypothetical protein
				SMD_0869	tig	cell division trigger factor
				SMD_0870	clpP	ATP-dependent Clp protease proteolytic subunit
				SMD_0871	clpX	ATP-dependent Clp protease ATP-binding subunit ClpX
				SMD_0872	lon	ATP-dependent protease La
				SMD_0873	hupB	DNA-binding protein HU-alpha
				SMD_0879	pPAID	peptidyl-prolyl cis-trans isomerase pPAID
				SMD_0880	Int	integrase
				SMD_0881		hypothetical protein

GEI	Start	End	Size	Locus ID	Gene name	Product
				SMD_0882		DNA-binding protein
				SMD_0883		hypothetical protein
				SMD_0884		hypothetical protein
				SMD_0885		hypothetical protein
				SMD_0886		hypothetical protein
				SMD_0887		hypothetical protein
				SMD_0888		hypothetical protein
				SMD_0889		hypothetical protein
				SMD_0890		plasmid stabilization protein
				SMD_0891		hypothetical protein
				SMD_0892	y4eB	y4eB gene in pNGR234a
				SMD_0893	pbsX	transcriptional regulator, PbsX family
				SMD_0894		lipoprotein
				SMD_0895		hypothetical protein
				SMD_0896		hypothetical protein
				SMD_0897		Plasmid replication initiator protein
				SMD_0898		chromosome (plasmid) partitioning protein ParA
				SMD_0899		hypothetical protein
				SMD_0900		glycosidases
				SMD_0901	TraF/T4SS	type IV secretory pathway, protease TraF
				SMD_0902	virD2/T4SS	type IV secretory pathway, VirD2 components (relaxase)
				SMD_0903		sensor signal transduction histidine kinase
				SMD_0904	phoB2	winged helix family two component transcriptional regulator
				SMD_0905	PE-PGRS	PE-PGRS FAMILY PROTEIN
				SMD_0906		hypothetical protein
				SMD_0907		hypothetical protein
				SMD_0908		hypothetical protein
GEI-6	1027006	1031822	4816	SMD_0912		RND efflux membrane fusion protein
				SMD_0913		Two component transcriptional regulator, winged helix family
				SMD_0914		sensory histidine kinase in two-component regulatory system with OmpR
				SMD_0915		hypothetical protein
				SMD_0916		Putative outer membrane protein CC_0351 precursor
GEI-7	1031382	1042897	11515	SMD_0916	CC_0351	outer membrane protein CC_0351 precursor
				SMD_0918	LysR-reg	LysR family transcriptional regulator
				SMD_0919		lipoprotein
				SMD_0920	virD4	type IV secretion system protein VirD4
				SMD_0921	copG	CopG domain-containing protein
				SMD_0922	trbB	Conjugative transfer protein TrbB
				SMD_0923	trbC	Conjugative transfer protein TrbC
				SMD_0924	trbD	Conjugative transfer protein TrbD
				SMD_0925	trbE	Conjugative transfer protein TrbE
				SMD_0926	trbJ	Conjugative transfer protein TrbJ
				SMD_0927		lipoprotein
				SMD_0928	trbL	Conjugative transfer protein TrbL
GEI-8	1055153	1074404	19251	SMD_0941		ADP-heptose--lipooligosaccharide heptosyltransferase II
				SMD_0942		3-deoxy-D-manno-octulosonic acid kinase
				SMD_0943		hypothetical protein
				SMD_0944		hypothetical protein
				SMD_0946	sugE	quaternary ammonium compound-resistance protein sugE
				SMD_0947		autotransporter protein
				SMD_0948		hypothetical protein
				SMD_0949		hypothetical protein
				SMD_0950	tadC/flpP	type II/IV secretion system protein TadC,associated with Flp PAIlus assembly
				SMD_0951	T2SS prot	type II secretion system protein
				SMD_0952	tadA/cpaF	type II/IV secretion system ATP hydrolase TadA/VirB11/CpaF, TadA subfamily
GEI-9	1549512	1558514	9002	SMD_1382	pstA	permease protein PstA (TC 3.A.1.7.1)
				SMD_1383	pstC	permease protein PstC (TC 3.A.1.7.1)
				SMD_1384	pstS	phosphate ABC transporter, (TC 3.A.1.7.1)
				SMD_1385	phoX	phosphate ABC transporter, (TC 3.A.1.7.1)
				SMD_1386		hypothetical protein
				SMD_1387	nth	endonuclease III
				SMD_1388		hypothetical protein
				SMD_1389	crt2	3-hydroxybutyryl-CoA dehydratase
				SMD_1390		macrophage infectivity potentiator
GEI-10	1648522	1669064	19104	SMD_1476	smmD2	Heavy metal RND efflux outer membrane protein,CzcC family
				SMD_1477	ISD2	ISStmaD2 Transposase A
				SMD_1478	ISD2	ISStmaD2 Transposase B

GEI	Start	End	Size	Locus ID	Gene name	Product
				SMD_1479	ISD1	ISStmaD1 Transposase B
				SMD_1480	ISD1	ISStmaD1 Transposase A
				SMD_1481	ISD5	ISStmaD5 transposase A
				SMD_1482	ISD5	ISStmaD5 transposase B
				SMD_1483	LysR	LysR family transcriptional regulator YnfL
				SMD_1484		hypothetical protein
				SMD_1485	copD	Copper resistance protein D
				SMD_1486	copC	Copper resistance protein C precursor
				SMD_1487	copF	Lead, cadmium, zinc and mercury transporting ATPase
				SMD_1488	Asl7591	Asl7591 protein
				SMD_1489	copG	copper amine oxidase N-terminal protein
				SMD_1490	copB	Copper resistance protein B
				SMD_1491	copA	Multicopper oxidase
				SMD_1493		Heavy metal sensor histidine kinase
				SMD_1494	arsR2	ArsR family transcriptional regulator
				SMD_1495		Lactoylglutathione lyase
				SMD_1496	arsC3	Arsenate reductase
				SMD_1497	acr3	arsenical-resistance protein ACR3
GEI-11	1670959	1676372	5413	SMD_1500	chrA	chromate transport protein ChrA
				SMD_1501	Acetyl-CoA	acetyltransferase
				SMD_1502		plasmid stabilization protein
				SMD_1504	smeV2	RND multidrug efflux membrane fusion protein
				SMD_1505	dlat	dihydrolipoamide acyltransferase (E2)
GEI-12	1693621	1698278	4657	SMD_1524		hypothetical protein
				SMD_1525		candidate type III effector Hop protein
				SMD_1526		excisionase domain-containing protein
				SMD_1527		hypothetical protein
				SMD_1528	ATPase	AAA ATPase
				SMD_1529		hypothetical protein
				SMD_1530		hypothetical protein
				SMD_1531		hypothetical protein
GEI-13	1710319	1739152	28833	SMD_1542		hypothetical protein
				SMD_1543		hypothetical protein
				SMD_1544		Plasmid related protein
				SMD_1545		hypothetical protein
				SMD_1546		hypothetical protein
				SMD_1547		hypothetical protein
				SMD_1548		Plasmid-related protein
				SMD_1549		hypothetical protein
				SMD_1550		hypothetical protein
				SMD_1551	tonB	periplasmic protein TonB, links inner and outer membranes
				SMD_1552		hypothetical protein
				SMD_1553		hypothetical protein
				SMD_1555	merA	mercuric ion reductase
				SMD_1556	merP	periplasmic mercury(2) binding protein
				SMD_1557	merT	mercuric transport protein, MerT
				SMD_1558	merR	mercuric resistance operon regulatory protein
				SMD_1559		hypothetical protein
				SMD_1560		hypothetical protein
				SMD_1561		hypothetical protein
				SMD_1562		hypothetical protein
				SMD_1563		hypothetical protein
				SMD_1564		hypothetical protein
				SMD_1565	cadA	Lead, cadmium, zinc and mercury transporting ATPase
				SMD_1566	czcD	Cobalt-zinc-cadmium resistance protein CzcD
				SMD_1567	topB	DNA topoisomerase III
				SMD_1568	PFGI-1 like	Single-stranded DNA-binding protein in PFGI-1-like cluster
				SMD_1569	inrR	integrase regulator R
				SMD_1570		hypothetical protein
				SMD_1571		hypothetical protein
GEI-14	1761475	1774086	12611	SMD_1598		hypothetical protein
				SMD_1599		hypothetical protein
				SMD_1600		hypothetical protein
				SMD_1601		hypothetical protein
				SMD_1602		hypothetical protein
				SMD_1603		hypothetical protein
				SMD_1604		Phage terminase, large subunit
				SMD_1605		hypothetical protein
				SMD_1606	B	Phage capsid and scaffold
				SMD_1607	clpP	peptidase S14, ClpP

GEI	Start	End	Size	Locus ID	Gene name	Product
GEI-15	1779635	1790559	10924	SMD_1614		hypothetical protein
				SMD_1615		hypothetical protein
				SMD_1616		hypothetical protein
				SMD_1617		hypothetical protein
GEI-16	1955157	1966250	11093	SMD_1778		hypothetical protein
				SMD_1779	ISD1-TranspA	ISStmaD1 Transposase A
				SMD_1780	ISD1-TranspB	ISStmaD1 Transposase B
				SMD_1781	Int	Integrase
GEI-17	2060200	2069193	8993	SMD_1862	hsdR	Type I restriction-modification system,restriction subunit R
				SMD_1863	hsdM	Type I restriction-modification system,DNA-methyltransferase subunit M
				SMD_1864	hsdS	Type I restriction-modification system,specificity subunit S
				SMD_1865		ABC transporter
				SMD_1866		hypothetical protein
GEI-18	2424072	2437217	13145	SMD_2174		hypothetical protein
				SMD_2175		Zona occludens toxin
				SMD_2176		hypothetical protein
				SMD_2177		DNA mismatch repair
				SMD_2178	tRNA-Pseudo	tRNA-OTHER
				SMD_2179	C4.5	c4 antisense RNA (Rfam family RF01695)
				SMD_2180		hypothetical protein
				SMD_2181		CTP:molybdopterin cytidylyltransferase
				SMD_2182	xdhC	xanthine dehydrogenase accessory factor
GEI-19	2462648	2472668	10020	SMD_2210		BLUF domain containing protein
				SMD_2211		BLUF domain containing protein
				SMD_2212		Glycosyl transferase, group 2 family protein
				SMD_2213		type 12 methyltransferase
				SMD_2214		LmbE-like protein
				SMD_2215		Acyl-CoA dehydrogenase/oxidase domain protein
				SMD_2216		hypothetical protein
				SMD_2217	ybdR	zinc-type alcohol dehydrogenase-like protein ybdR
				SMD_2218		hypothetical protein
				SMD_2219		hypothetical protein
				SMD_2220		isochorismatase
				SMD_2221		hypothetical protein
				SMD_2222	Tiorf68	Tiorf68 protein
				SMD_2223		hypothetical protein
				SMD_2224	lacL	Transcriptional regulator lacI family
GEI-20	2509919	2517701	7782	SMD_2247		Na /H antiporter NhaD and related arsenite permeases
				SMD_2248	OprO/OprP	phosphate-specific outer membrane OprP/OprO
				SMD_2249		ABC transporter permease
				SMD_2250		ABC transporter ATP-binding protein
				SMD_2251		ABC transporter substrate-binding protein
				SMD_2252		hypothetical protein
				SMD_2253	ybhD	LysR family transcriptional regulator YbhD
				SMD_2254		hypothetical protein
				SMD_2255		hypothetical protein
GEI-21	2962972	2969901	6929	SMD_2661		hypothetical protein
				SMD_2662		putative signal transduction protein with EFhand domain
				SMD_2663		Hemolysin related protein
				SMD_2664		hypothetical protein
				SMD_2665		Hemolysin related protein
				SMD_2666		hypothetical protein
				SMD_2667		hypothetical protein
				SMD_2668	C4.6	c4 antisense RNA (Rfam family RF01695)
				SMD_2669		HxlR family transcriptional regulator
GEI-22	4197405	4233316	35911	SMD_3750	ygiF	inner membrane protein YqjF
				SMD_3751		hypothetical protein
				SMD_3752		hypothetical protein
				SMD_3753		Secreted and surface protein containing fasciclin-like repeats
				SMD_3754		hypothetical protein
				SMD_3755		hypothetical protein
				SMD_3756	yjogF	translation initiation inhibitor,yjgFfamily
				SMD_3757		hypothetical protein
				SMD_3758		hypothetical protein
				SMD_3759		hypothetical protein
				SMD_3760	ISD1-	ISStmaD1 Transposase B

GEI	Start	End	Size	Locus ID	Gene name	Product
					TranspB ISD1-	
				SMD_3761	TranspA	ISStmaD1 Transposase A
				SMD_3762	Int	integrase
				SMD_3764		hypothetical protein
				SMD_3765	rpoD	RNA polymerase sigma factor RpoD
				SMD_3766	dtd	D-tyrosyl-tRNA(Tyr) deacylase
GEI-23	4430040	4436910	6870	SMD_3959		rifamPAIn ADP-ribosyl transferase
				SMD_3960		hypothetical protein
				SMD_3961		hypothetical protein
				SMD_3962	C4.7	c4 antisense RNA (Rfam family RF01695)
				SMD_3963		hypothetical protein
				SMD_3964		hypothetical protein

Appendix 4: Putative virulence genes matrix generated analyzing the 24 genomes of *Stenotrophomonas maltophilia* presented in this work. For all genes, the sequences of D457 were used as reference. When another reference sequence was used, it is mention as follow: (*) sequence obtained from *S. maltophilia* K279a genome; (**) sequence obtained from *S. maltophilia* RA8 genome; (***) sequence obtained from *S. maltophilia* St53; (****) sequence obtained from *S. maltophilia* SKK35 strain. Gray cells indicate the presence of the gene; white (empty) cells indicate the absence.

Reference sequence	Putative virulence genes of *S. maltophilia*	Gene	K279a	D457	JV3	R551-3	FL1	FL2	FL3	FL4	FL5	FL6	FL7	FL8	FL9	FL10	FL11	FL12	FL13	FL14	FL15	FL16	FL17	FL18	FL19	FL20
GI:504460615	DNase																									
GI:504460964	Hemolysin III																									
GI:754362796	Non fimbrial adhesin	afaD																								
GI:765003484*	Filamentous hemagglutinin																									
GI:765004840*	Filamentous hemagglutinin																									
GI:493444609**	Filamentous hemagglutinin	fhaB																								
GI:765004612*	Hemagglutinin like adhesin																									
GI:504460261*	T5SS autotransporter hemagglutinin_1																									
GI:754362361	T5SS autotransporter hemagglutinin_2																									
GI:504461648	Hemoglobin binding protein	hgbB																								
GI:504460839	Hemoglobin binding protein	hgbC																								
GI:754361938	Serine-protease	StmPr1																								
GI:504458550	Serine-protease																									
GI:504458785	Serine-protease																									
GI:504458833	Serine-protease																									
GI:504459284	T5SS serine-protease																									
GI:754362539	Serine-protease	sphB																								
GI:504461425	T5SS serine-protease																									
GI:504461644	Serine-protease																									
GI:493412491**	Serine-protease																									
GI:504461623	Serine-protease	StmPr3																								
GI:754361966	Serine-protease	yapH																								
GI:754362756	Protease IV	sppA																								
GI:504458441	Metalloprotease	ppqL/ptrA																								
GI:754361759	Metalloprotease																									
GI:754361869	Metalloprotease																									
GI:504459393	Metalloprotease																									
GI:504461110	Autotransporter lipase/esterase																									

Reference sequence	Putative virulence genes of *S. maltophilia*	Gene	K279a	D457	JV3	R551-3	FL1	FL2	FL3	FL4	FL5	FL6	FL7	FL8	FL9	FL10	FL11	FL12	FL13	FL14	FL15	FL16	FL17	FL18	FL19	FL20
GI:501455617*	Putative Lipase																									
GI:696384648	Phospholipase B_1																									
GI:501457238*	Phospholipase B_2																									
GI:504459627	Phospholipase C	plcN1																								
GI:504460620	Phospholipase A																									
GI:754362212	Phospholipase D_1																									
GI:504461767	Phospholipase D_2																									
GI:504458309	Phospholipase D_3																									
GI:754362539	Phospholipase D_4																									
GI:504460314	Siderophore enterobactin synthetase_1	EntACF_1																								
GI:504460315	Siderophore enterobactin synthetase_2	EntACF_2																								
GI:504460319	Siderophore enterobactin synthetase_3	EntACF_3																								
GI:692316813***	RTX-toxin-activating protein																									
GI:493496316****		FrpA/C																								
GI:493491520****		lktD																								
GI:493491522****		hlyB																								
GI:491555445		hlyD																								

Appendix 5: Analysis of the variance for the susceptibility tests and MICs found to the different groups with a p-value of 0,05.

SUMMARY					ANOVA						
Groups	*Count*	*Sum*	*Arith. Mean*	*Variance*	*Source of variations*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit.*
SXTc	10	5.22	52%	0.0657	Between Groups	0.54615	1	0.54615125	11.33	0.00	4.41
SXTe	10	1.915	19%	0.0305	Within Groups	0.86694	18	0.048163694			
					Total	1.41309	19				
TGCc	10	16.594	1.6594	3.1080	Between Groups	7.93422	1	7.93422045	4.98	0.03	4.41
TGCe	10	3.997	0.3997	0.0730	Within Groups	28.63006	18	1.590558917			
					Total	36.56428	19				
CAZc	10	1202.5	120.25	12737.180	Between Groups	23839.5125	1	23839.5125	2.63	0.1221	4.41
CAZe	10	512	51.2	5378.8444	Within Groups	163044.225	18	9058.0125			
					Total	186883.7375	19				
PMc	10	630	63	5567.333	Between Groups	168.2	1	168.2	0.03	0.85	4.41
PMe	10	688	68.8	4497.066	Within Groups	90579.6	18	5032.2			
					Total	90747.8	19				
CNc	10	102.88	10.288	153.7738	Between Groups	243.50428	1	243.5042898	3.08	0.09	4.41
CNe	10	33.094	3.3094	3.943061	Within Groups	1419.45252	18	78.85847291			
					Total	1662.95682	19				
GATc	10	132.861	13.2861	1625.436	Between Groups	179.17893	1	179.1789384	0.10	0.74	4.41
GATe	10	192.724	19.2724	1635.315	Within Groups	29346.76239	18	1630.375688			
					Total	29525.94132	19				
CSc	10	417	41.7	5870.233	Between Groups	2184.05	1	2184.05	0.71	0.40	4.41
CSe	10	208	20.8	255.2888	Within Groups	55129.7	18	3062.761111			
					Total	57313.75	19				
CLc	10	237	23.7	1430.677	Between Groups	572.45	1	572.45	0.15	0.70	4.4139
CLe	10	344	34.4	6082.488	Within Groups	67618.5	18	3756.583333			
					Total	68190.95	19				
IMIc	10	320	32	0	Between Groups	50.99540	1	50.9954048	1.00	0.33	4.41
IMIe	10	288.064	28.8064	101.9908	Within Groups	917.91728	18	50.9954048			
					Total	968.91269	19				
ETPc	10	320	32	0	Between Groups	50.593805	1	50.593805	1.00	0.33	4.41
ETPe	10	288.19	28.819	101.18761	Within Groups	910.68849	18	50.593805			
					Total	961.282295	19				
MXFc	10	4.787	0.4787	0.8259406	Between Groups	0.86320	1	0.86320125	2.08	0.16	4.41
MXFe	10	0.632	0.0632	0.0014950	Within Groups	7.44692	18	0.413717872			
					Total	8.310122	19				
NAc	10	97	9.7	184.01111	Between Groups	212.87812	1	212.878125	2.25	0.15	6.51
NAe	10	31.75	3.175	4.7506944	Within Groups	1698.85625	18	94.38090278			
					Total	1911.73437	19				

Appendix 6: List of resistance genes and the number of bireads recognized in each sample at the Experiment I. Resistance genes into the gray cells represent the genes found in plasmids and the resistance genes in white cells represent the genes found usually in chromosomes.

Resistance genes	ARPCARD Description	Control I	100 ppt	100 ppb	500 ppb	1 ppm	10 ppm
NC_002156.4594907	POX7-1_p1.	31184	393	398959	8354	15342	1539
HQ451074.1.gene4	blaTEM-1.	15148	979	198903	19838	30931	816
AF332513.1.gene1	blaTEM-63.	10973	2083	133746	34339	54769	544
AF091113.2.gene1	TEM-67.	10170	1787	129945	29544	44572	444
NC_011586.7045197	AB57_0283.	6198	89	88917	2033	3210	340
AY589493.1.gene1	blaTEM-112.	15	22	8	3	144	327
FJ919776.1.gene1	blaTEM-168.	5095	54	65850	1222	2393	275
M37699.gene1	tetX.	8		39	58	50	163
AF188199.1.gene1	blaTEM-70.	8	37	43	56	121	72
AJ318093.1.gene1	bla-TEM-93.	1	12	7	31	197	67
AY039040.1.gene1	TEM-89.	1	8	2	6	80	51
AY027590.1.gene1	AAK14792.1.	2	1	2	1	34	37
AB049569.1.gene1	TEM-91.	3	29	11	33	64	30
DQ865198.gene	aadA5.			5	3	1	25
AF155139.2.orf0.gene	vanRF.						21
AJ704863.gene12	aadA1.	7	28	6	27	72	18
DQ286729.1.gene1	ABB97007.1.	3	5	1	9	58	13
AY123251.gene3	aadA1.	3	5	1	11	5	13
AF397068.1.gene1	AAK85245.1.	1	14	1	1	5	12
AF351241.1.gene1	blaTEM-90.	6	2	3	17	13	11
AY289608.1.gene2	oxa-53.						11
AF506748.1.gene1	AAM28884.1.	11	2	6	1	1	10
AF495873.1.gene1	AAM18924.1.	4	6	3	9	72	7
AM941159.1.gene1	blaTEM.	4	5	10	17	53	7
AY264780.2.gene3	tetA(41).		7	3	11	106	6
AF190692.1.gene1	blaTEM-79.	5	12	2	6	10	6
AJ239002.1.gene1	bla-TEM-68.	3	10	1		2	6
AY628199.1.gene1	blaTEM-126.		7	1	4	23	5
AY628176.1.gene1	blaTEM-125.	1	3	1	4	23	5
AM087454.1.gene1	blaTEM.	1	10			3	5
NC_010558.1.6275948	QepA.			1		2	5
AY307100.1.gene1	AAQ98890.1.	4	1	7	9	30	4
JF268688.1.gene6	catB.	6		8	1	5	4
EU118119.1.orf1.gene	aadA2.	1	2	2	7	16	3
HQ451074.1.gene20	sul1.	3		39	7	1	3
EF468463.1.gene1	bla.	3	1	46	270	357	2
AY458224.gene	sul1**.	17	11	2	45	106	2
AF397067.1.gene1	AAK85244.1.			23	38	50	2
GU550123.1.gene1	ADB90239.1.	9	2	86	24	50	2
AY103455.gene	aac(6)-Ib*.	1	7	2		50	2
AJ420864.1.gene1	blaIMP.	9	582	19	105	44	2
FJ807656.1.gene1	blaTEM-154.	2		23	43	27	2
DQ464881.1.gene2	sul2.					18	2
AF516719.1.gene1	AAM61952.1.			15	14	12	2
EF136377.1.gene1	blaTEM-160.			5	1	8	2
EF636461.1.orf1.gene	aac(6')-Ib-cr.		6		7	6	2
AM183304.1.gene1	blaTEM-150.	2	1	6	5	4	2
AY139598.1.gene3	oxa.			8	17	2	2
NC_011586.7045194	cat.			1			2
JN211012.1.gene1	blaTEM-188.	1	1				2

Resistance genes	ARPCARD Description	Control I	100 ppt	100 ppb	500 ppb	1 ppm	10 ppm
AJ222769.gene	tetW*.	8	5	93	70	132	1
EU274580.1.gene1	ABX71157.1.	13	2	113	46	95	1
AY092401.1.gene1	AAM22276.1.	21	3	225	59	78	1
AJ550807.1.gene4	bla-imp13.					70	1
AF397066.1.gene1	AAK85243.1.	4	8		7	52	1
AF190693.1.gene1	blaTEM-78.			30	33	50	1
HQ529916.1.gene1	TEM-183.	20	1	187	10	44	1
AY101764.1.gene1	blaTEM-107.					37	1
EF118171.1.gene7	sul1.	1	3		14	36	1
DQ105528.2.gene1	TEM-145.	31	1	312	22	32	1
AY743590.gene	tet39.			6	20	22	1
AF319779.2.orf0.gene	erm(35).	1	19	1	35	17	1
EF534736.1.gene1	blaTEM-158.	4		70	13	17	1
AF190694.1.gene1	blaTEM-76.			29	35	8	1
HQ451074.1.gene24	armA.p01.	1	1		8	4	1
HM246246.1.gene1	ADM61585.1.	6		49	1	3	1
L20800.gene	tetBP.	1		36	1	3	1
AY327540.1.gene1	blaTEM-124.	1	7	1		3	1
AY368237.1.gene1	AAR89359.1.	1	4			3	1
FJ405211.1.gene1	TEM-57.	3		81	1	1	1
NC_011595.7057747	ABBFA_002299.					1	1
X04388.gene	tetM.						1
X75439.1.gene2	mupA.						1
DQ212986.1.gene4	vanRG.				9	217	
AM849805.1.gene1	blaTEM-15.	8		81	93	125	
Y14574.2.gene1	bla	9		87	14	107	
AF253562.2.orf5.gene	vanYG2.					84	
JN254627.1.gene1	tem-189.	9	1	131	52	67	
AF427129.1.gene1	blaTEM-83.			22		60	
AJ420072.gene	tet33.			10	34	51	
AF104442.1.gene1	blaTEM-54.	10	1	128	17	51	
AJ866988.1.gene1	tem-130.					48	
AY553332.1.gene2	aac(6').					35	
NC_010410.6002874	oxa-10.	2		2	35	33	
AY740681.1.gene5	catB3.					32	
EU815939.1.gene1	ACF32746.1.	5		56	13	28	
AY628175.1.gene1	blaTEM-109.	6	1	69	15	27	
DQ105529.2.gene1	TEM-146.	15		218	10	27	
AF427127.1.gene1	blaTEM-81.	5	1	75	18	26	
AY605049.2.gene1	blaVIM-11.	2		69	10	24	
D85892.1.orf0.gene	mphB.					21	
AJ437107.1.gene1	bla.	8		101	1	19	
FJ360884.1.gene1	blaTEM-167.					18	
AY436361.1.gene1	blaTEM-131.	5		113	11	17	
AY327539.1.gene1	blaTEM-123.			7	4	17	
DQ294299.gene	tetW_2.	6		58	4	16	
GU208678.1.gene1	blaGES-15.					16	
JN416112.1.gene1	tem-190.	12		116	11	15	
GU371926.1.gene99	blaCTX-M-15.			4	105	14	
HQ874631.1.gene1	ADZ48685.1.					14	
AY183453.1.orf0.gene	ereA.		8		7	13	
DQ279850.1.gene1	ABB84515.1.			5	4	13	
NC_005054.2598277	vanR.			1	2	10	

Resistance genes	ARPCARD Description	Control I	100 ppt	100 ppb	500 ppb	1 ppm	10 ppm
FN652295.1.gene1	blaTEM-177.	12	1	109	5	9	
GU371926.1.gene95	blaTEM-33.	1	1	4	4	9	
AY589494.1.gene1	blaTEM-113.	6	1	75	2	8	
NC_002516.2.878374	cat.		5	1	2	8	
AF468003.1.gene1	AAL77062.1.			19	26	7	
AY826417.1.gene1	blaTEM-136.	5		64	6	7	
AF427128.1.gene1	blaTEM-82.	5	1	34	4	7	
DQ834728.1.gene1	ABI74447.1.			5	3	7	
AF133139.gene	tetG.	4	10	1	2	5	
JF949915.1.gene1	blaTEM-1.	1	1		1	5	
AF518567.2.gene4	blactx-m-25.			12	2	4	
AM286274.1.gene1	blaTEM-137.			9		4	
DQ909059.1.gene1	ABI81768.1.	12	1	139	7	3	
M95287.gene	blaOXA-2.	1		27	3	3	
AY271267.1.gene1	blaTEM-121.	8		71	2	3	
NC_010410.6003170	aadA1.			1		3	
NC_010410.6002585	aacC1.		1			3	
NC_010481.6155789	blaOXA-58.					3	
AF516720.1.gene1	AAM61953.1.			12	4	2	
AJ634602.gene	blaTEM.			9	4	2	
DQ369751.1.gene1	blaTEM-149.			5	4	2	
FJ197316.1.gene1	ACI25375.1.	1	1	41	2	2	
AJ308558.1.gene1	blaTEM-95b.			12		2	
FJ873740.1.gene1	TEM.			5		2	
NC_011586.7045516	AB57_0437.					2	
NC_011586.7045804	AB57_2380.					2	
JF268688.1.gene5	blaOXA-209.			4	8	1	
DQ834729.1.gene1	ABI74448.1.	2		53	6	1	
NC_011595.7057349	ABBFA_000784.				1	1	
AJ584652.2.gene7	aadA1.			10		1	
NC_011595.7057325	ABBFA_000209.		1			1	
NC_011595.7057907	ABBFA_000816.					1	
NC_011595.7059276	ABBFA_002430.					1	
U82965.2.orf14.gene	U82965.2.orf14.					1	
AJ632249.1.gene1	oxa-59.				12		
JN227084.1.gene1	blaTEM-186.			4	1		
AY368236.1.gene1	AAR89358.1.			7			
DQ075245.1.gene1	blaTEM-143.			6			
NC_010410.6003168	aadA1.	2		1			
V01547.1.orf0.gene	V01547.1.orf0.			1			
X01702.1.orf2.gene	X01702.1.orf2.			1			
Y10281.1.gene1	bla			1			
Y17583.1.gene1	blaTEM-22.			1			
Y18050.2.gene6	aadA1.			1			
Y19114.gene	TetC.			1			
AJ584652.2.gene5	aac(6')-30/aac(6')-Ib'.		18				
AM932669.1.gene4	aadA1.p01.		17				
U37105.2.gene4	aadA10.		1				
U57969.gene	mexD.		1				
U59183.1.gene2	aac(6')-Ib.		1				
X97254.1.gene1	blaZ		1				
Y16952.3.orf35.gene	vanR.		1				

Resistance genes	ARPCARD Description	Control I	100 ppt	100 ppb	500 ppb	1 ppm	10 ppm
NC_010410.6002612	tetA.	4					
NC_008702.1.4609454	macB.	1				53	56
NC_002516.2.881071	mexD.		2	1	2	9	37
NC_011586.7046391	adeA.					1	23
NC_008702.1.4607352	msbA1.	1		2		37	5
NC_010400.5984386	ABSDF0738.					5	5
NC_008702.1.4607362	azo2334.	1		1	1		3
NC_002516.2.882884	mexF.	6		1			3
NC_002695.1.916585	ECs2883.	2		13	119	102	2
NC_008702.1.4606597	azo0833.		2			3	2
NC_008702.1.4607484	azo2569.	4		20	65	33	1
NC_008702.1.4607729	azo2993.		1	20	14	14	1
NC_002516.2.877851	oprM.					13	1
NC_002695.1.917702	ECs0964.	2		6	5	11	1
NC_010410.6000769	ABAYE0746.					3	1
M14730.gene	ermF*.	7	1	2			1
NC_011595.7058895	msbA.		1				1
NC_002695.1.914737	ECs3547.	2		17	46	78	
AJ579365.gene	lsaB_orf3.					53	
AL939114.1.orf1.gene	SCO2860.					40	
DQ072853.1.gene1	bla-TEM-139.			6	5	34	
NC_002758.1121879	SAV1866.	1		4	13	26	
NC_008702.1.4608898	azo0443.			1	1	26	
NC_008702.1.4606330	blaA.				1	24	
NC_002695.1.914620	ECs0516.	1		7	9	22	
NC_010079.5775045	USA300HOU_0045.			1		22	
NC_002695.1.915267	ECs3332.			2	8	16	
NC_002695.1.915420	emrD.			2	7	16	
NC_002695.1.912777	ECs1864.	4		53	9	14	
NC_002516.2.880417	mexT.					13	
NC_002516.2.880556	PA5294.					12	
AY956335.1.gene1	AAX56615.1.	8		55	3	11	
NC_002516.2.882885	oprN.					11	
FM897214.1.orf2.gene	arr.	3		46	8	10	
NC_010400.5984385	ABSDF0737.			1		10	
NC_008702.1.4606435	rnfE1.					10	
NC_002695.1.914764	ECs3520.			2	9	9	
NC_008702.1.4607719	cphA.					8	
NC_008702.1.4606607	azo0847.	1		2	6	6	
NC_002516.2.877852	mexB.	1		26	3	6	
NC_009085.4919117	A1S_1750.				1	6	
NC_010400.5984384	ABSDF0736.					6	
NC_002695.1.913273	emrE.	1		10	3	5	
NC_002695.1.914619	ECs0515.	3		21		5	
NC_002695.1.916584	ECs2882.			2		5	
NC_010400.5985985	macB.					5	
NC_002695.1.916586	ECs2884.			2		4	
NC_002695.1.914736	ECs3548.			2	34	3	
NC_002695.1.914045	ampC.	1		9	14	3	
NC_008702.1.4606278	oprM1.		2		1	3	
NC_010410.6003177	emrA.					3	
NC_010410.6003262	emrB.					3	
NC_008702.1.4607810	azo3134.		1	11	6	2	

Resistance genes	ARPCARD Description	Control I	100 ppt	100 ppb	500 ppb	1 ppm	10 ppm
M15332.gene	erm(G)**_ermG*.				6	2	
NC_008702.1.4608011	msbA2.		1	6	2	2	
NC_002516.2.880346	mexI.		3		1	2	
NC_002695.1.916587	ECs2885.			2		2	
NC_008702.1.4607363	azo2335.	1		1		2	
NC_011586.7045443	adeI.					2	
NC_011586.7045550	abeM.					2	
NC_002695.1.915653	ECs3247.		2	1	1	1	
NC_011586.7045444	adeJ.				1	1	
NC_002695.1.915750	ECs4393.	2		17		1	
NC_011595.7058890	ABBFA_003020.					1	
NC_011595.7059574	ABBFA_001708.					1	
NC_011595.7059713	ABBFA_000731.					1	
NC_002655.gene	msr(A).				6		
NC_002516.2.881672	msbA.	1		18	1		
NC_011595.7060505	ABBFA_003019.				1		
NC_008702.1.4606277	azo0245.	5		1			
NC_008702.1.4607935	azo3375.	1		1			
NC_011595.7058613	ABBFA_003018.			1			
NC_012469.1.7685970	SPT_1925.			1			
NC_012469.1.7686946	SPT_0145.			1			
NC_008702.1.4606598	azo0834.		2				
NC_011595.7058956	ABBFA_000034.	8					
NC_008702.1.4606564	azo0784.	3					
NC_010685.6295744	pKH19_p2.	3					
NC_012469.1.7686878	SPT_1414.	1					

Appendix 7: List of resistance genes and the number of bireads recognized in each sample at the Experiment II. Resistance genes into the gray cells represent the genes found in plasmids and the resistance genes in white cells represent the genes found usually in chromosomes.

Resistance genes	ARPCARD Description	Control II	1 ppm	10 ppm
NC_002156.4594907	POX7-1_p1	42542	998	257
HQ451074.1.gene4	blaTEM-1	22837	2405	143
AF332513.1.gene1	blaTEM-63	14408	4987	90
AF091113.2.gene1	TEM-67	12317	3989	87
NC_011586.7045197	AB57_0283	9191	214	51
FJ919776.1.gene1	blaTEM-168	7013	158	57
DQ464881.1.gene2	sul2	50		2
HQ451074.1.gene20	sul1	43		8
DQ105528.2.gene1	TEM-145	34	3	
EU274580.1.gene1	ABX71157.1	24	4	
HQ529916.1.gene1	TEM-183	22	6	
AY589493.1.gene1	blaTEM-112	20		1
DQ105529.2.gene1	TEM-146	19	1	
AF427127.1.gene1	blaTEM-81	17		
AF104442.1.gene1	blaTEM-54	16	1	1
AY743590.gene	tet39	16		
GU550123.1.gene1	ADB90239.1	15	7	
JN254627.1.gene1	tem-189	15	5	
AY092401.1.gene1	AAM22276.1	14	7	
DQ909059.1.gene1	ABI81768.1	14		
JN416112.1.gene1	tem-190	14		
AM849805.1.gene1	blaTEM-15	13	4	
AY436361.1.gene1	blaTEM-131	13		
FN652295.1.gene1	blaTEM-177	13		
AF133139.gene	tetG	12	1	7
AJ437107.1.gene1	bla	12	2	
AF188199.1.gene1	blaTEM-70	12		
AF506748.1.gene1	AAM28884.1	11	2	
AM183304.1.gene1	blaTEM-150	11	2	
AB049569.1.gene1	TEM-91	11		
JF268688.1.gene6	catB	10		5
Y14574.2.gene1	bla	10	8	
AY271267.1.gene1	blaTEM-121	8	4	
AF190692.1.gene1	blaTEM-79	7	2	
AF427128.1.gene1	blaTEM-82	7	1	
HM246246.1.gene1	ADM61585.1	7	1	
AF397066.1.gene1	AAK85243.1	6		
EF534736.1.gene1	blaTEM-158	6		
AM087454.1.gene1	blaTEM	5		
AY307100.1.gene1	AAQ98890.1	5		
NC_010410.6002612	tetA	4		7
AJ239002.1.gene1	bla-TEM-68	4	1	1
AY628175.1.gene1	blaTEM-109	4		1
AY826417.1.gene1	blaTEM-136	4	3	
AF495873.1.gene1	AAM18924.1	4		
AM941159.1.gene1	blaTEM	4		
AY628176.1.gene1	blaTEM-125	4		
FJ807656.1.gene1	blaTEM-154	4		
GU371926.1.gene95	blaTEM-33	4		
JF949915.1.gene1	blaTEM-1	4		
NC_010481.6155789	blaOXA-58	4		

Resistance genes	ARPCARD Description	Control II	1 ppm	10 ppm
X97254.1.gene1	blaZ	3	3	
AF397068.1.gene1	AAK85245.1	3	2	
AY039040.1.gene1	TEM-89	3	1	
AJ318093.1.gene1	bla-TEM-93	3		
NC_011595.7057907	ABBFA_000816	3		
AJ704863.gene12	aadA1	2		5
AY589494.1.gene1	blaTEM-113	2	4	
DQ286729.1.gene1	ABB97007.1	2	1	
EU815939.1.gene1	ACF32746.1	2	1	
AF155139.2.orf0.gene	vanRF	2		
AF351241.1.gene1	blaTEM-90	2		
U57969.gene	mexD	1	1	8
AY458224.gene	sul1**	1	4	6
HQ451074.1.gene24	armA.p01	1	70	2
AF174129.3.gene10	blaCTX-M-9.p01	1		2
AJ420072.gene	tet33	1		1
AF468003.1.gene1	AAL77062.1	1	1	
AF516719.1.gene1	AAM61952.1	1	1	
AF516720.1.gene1	AAM61953.1	1	1	
AY327540.1.gene1	blaTEM-124	1	1	
AY027590.1.gene1	AAK14792.1	1		
AY327539.1.gene1	blaTEM-123	1		
DQ075245.1.gene1	blaTEM-143	1		
DQ369751.1.gene1	blaTEM-149	1		
EF136376.1.gene1	blaTEM-159	1		
FJ405211.1.gene1	TEM-57	1		
JN227084.1.gene1	blaTEM-186	1		
NC_011586.7045804	AB57_2380	1		
NC_011595.7059276	ABBFA_002430	1		
AB187515.1.orf3.gene	qnr			54
DQ865198.gene	aadA5			35
EU118119.1.orf1.gene	aadA2		2	24
AY139598.1.gene3	oxa			16
L42544.gene	tetT*			15
NC_010410.6002874	oxa-10			13
X04388.gene	tetM			12
AF518567.2.gene4	blactx-m-25			11
M95287.gene	blaOXA-2		1	9
AJ550807.1.gene4	bla-imp13			7
FN594949.1.gene24	tet(44)			7
JF800667.1.gene2	blaOXA-205		1	6
M37699.gene1	tetX			6
AF317511.1.gene5	blaOXA			6
AF371964.1.gene1	AAK55330.1			6
AY139598.1.gene2	aadA5			5
AY103455.gene	aac(6)-Ib*			4
AY123251.gene3	aadA1			4
GU371926.1.gene99	blaCTX-M-15			4
AY740681.1.gene5	catB3			3
L07945.1.gene1	blaOXA-3			3
Y19114.gene	TetC			3
AJ420864.1.gene1	blaIMP		3	2
AF319779.2.orf0.gene	erm(35)		1	2
X75439.1.gene2	mupA		1	2

Resistance genes	ARPCARD Description	Control II	1 ppm	10 ppm
AF357599.1.gene1	blaCMY-11			2
AJ584652.2.gene7	aadA1			2
AY289608.1.gene2	oxa-53			2
JF268688.1.gene5	blaOXA-209			2
Y19116.gene	tetE			2
AF030344.gene	tetV			1
AF255921.1.gene2	blaZ			1
AJ295238.gene	tet32			1
DQ294299.gene	tetW_2			1
EF118171.1.gene7	sul1			1
EF552405.1.gene1	blaOXA-141			1
FJ666065.1.gene1	ACQ82807.1			1
GQ152602.1.gene1	blaMOX-7			1
L06940.gene	TetE			1
NC_005054.2598277	vanR			1
NC_010410.6002875	veb-1			1
NC_010410.6003168	aadA1			1
U14748.1.gene2	blaAER-1			1
U37105.2.gene4	aadA1			1
Y10282.gene	blaFOX-2			1
Y16952.3.orf35.gene	vanR			1
Y17583.1.gene1	blaTEM-22		3	
AY628199.1.gene1	blaTEM-126		2	
U82965.2.orf14.gene	U82965.2.orf14		2	
AF190694.1.gene1	blaTEM-76		1	
EF468463.1.gene1	bla		1	
NC_002695.1.915267	ECs3332	163		
NC_002695.1.914619	ECs0515	141		2
NC_008702.1.4607362	azo2334	136		22
NC_002695.1.914045	ampC	83		
NC_002695.1.916585	ECs2883	48		14
NC_002695.1.917702	ECs0964	44		1
NC_002695.1.916586	ECs2884	42		11
NC_002695.1.915420	emrD	41		
NC_002695.1.913273	emrE	36		
NC_002695.1.916584	ECs2882	31		
NC_002695.1.914620	ECs0516	30		
NC_008702.1.4606277	azo0245	26	2	6
NC_002695.1.915750	ECs4393	24		2
NC_002695.1.914737	ECs3547	21		
NC_002695.1.916587	ECs2885	21		
NC_002655.gene	msr(A)	20		2
NC_002695.1.914764	ECs3520	19		2
NC_002695.1.914736	ECs3548	18		
NC_008702.1.4606564	azo0784	10	3	1
NC_008702.1.4609454	macB	9	13	9
NC_002695.1.915653	ECs3247	9		
NC_011595.7058956	ABBFA_000034	8		62
AY956335.1.gene1	AAX566151	7		
NC_011586.7045444	adeJ	7		
NC_002516.2.880346	mexI	6	1	14
M14730.gene	ermF*	6		5
NC_002516.2.877852	mexB	5	7	28
NC_011595.7059574	ABBFA_001708	5		

Resistance genes	ARPCARD Description	Control II	1 ppm	10 ppm
NC_008702.1.4607935	azo3375	4	4	1
NC_010410.6003262	emrB	4		
NC_011595.7059713	ABBFA_000731	4		
NC_011595.7058890	ABBFA_003020	3		
NC_002695.1.912777	ECs1864	2		67
NC_009085.4919117	A1S_1750	2	1	
NC_011595.7058895	msbA	2		
NC_011595.7060505	ABBFA_003019	2		
NC_002516.2.882884	mexF	1	8	54
NC_008702.1.4607484	azo2569	1	2	27
NC_008702.1.4607810	azo3134	1	2	11
NC_008702.1.4606607	azo0847	1	5	5
NC_002516.2.881071	mexD	1		2
NC_008702.1.4608011	msbA2	1	2	1
AE000516.2.gene3902	folP-2	1		
NC_010400.5984384	ABSDF0736	1		
NC_010410.6003177	emrA	1		
NC_011586.7045443	adeI	1		
NC_012469.1.7685970	SPT_1925			33
NC_002516.2.881672	msbA		1	17
NC_009782.5559087	ermA			8
NC_002516.2.880417	mexT			5
AF110130.1.orf0.gene	linB			4
NC_006663.1.orf0.gene	aadE			4
M15332.gene	erm(G)**_ermG*			3
NC_002516.2.877855	mexA			3
M64090.gene	ermT*			2
NC_002516.2.880212	mexE			2
NC_010400.5984909	ABSDF1462			2
NC_010400.5984910	ABSDF1463			2
NC_008702.1.4607352	msbA1		2	1
FN594949.1.gene22	blmA			1
M11189.1.gene1	AAA22276.1			1
NC_002516.2.882885	oprN			1
NC_008702.1.4608898	azo0443			1
AE000516.2.gene4101	embB.		1	
AJ579365.gene	lsaB_orf3		1	
NC_002516.2.877851	oprM.		1	
NC_008702.1.4606598	azo0834		1	

Publications